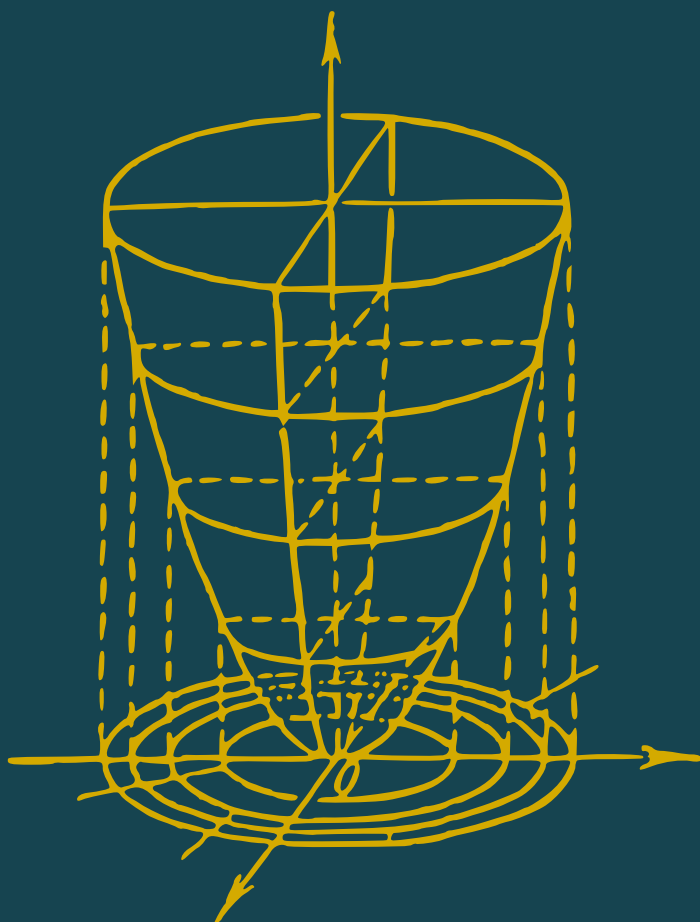


A.F. Bermant, I. G. Aramanovich

MATHEMATICAL ANALYSIS

*A Brief Course for
Engineering Students*



Mir Publishers Moscow

А. Ф. Бермант, И. Г. Араманович

КРАТКИЙ КУРС
МАТЕМАТИЧЕСКОГО АНАЛИЗА
ДЛЯ ВТУЗОВ

Издательство «Наука» Москва

*A. F. Bermant,
I. G. Aramanovich*

MATHEMATICAL ANALYSIS

*A Brief Course
for Engineering Students*

TRANSLATED FROM
THE RUSSIAN

by
V. M. VOLOSOV, D.Sc.
and I. G. VOLOSOVA

*Mir Publishers
Moscow*

First published 1975
Second edition 1979
Revised from the 1973 Russian edition
Third printing 1986

На английском языке

© Издательство «Наука», 1973

© English translation, Mir Publishers, 1975

Preface

This course is designed as a textbook for engineering students. It embraces the topics in mathematical analysis usually included into curricula of technical colleges. The course also contains some optional material which may be omitted in a first reading of the book; the corresponding items are marked with the asterisk.

There are a number of courses dealing with special divisions of mathematical analysis, such as equations of mathematical physics, functions of a complex argument and the like, and therefore, although these divisions are important for mathematical education of an engineer, they are not treated in this book. We also draw attention to the fact that only a few questions related to approximate calculations and programming (e.g. the application of the differential to approximate calculations, methods of approximate solution of equations, numerical integration and solution of differential equations, etc.) are discussed in this course. For a thorough study of this subject some other textbooks should be used.

In this course many examples are given which demonstrate the application of mathematical analysis to various divisions of mechanics and physics. The study of these examples is very important since the main interest of an engineer lies in solving concrete applied problems. At the end of each chapter we give a number of questions aimed at checking the understanding of the theoretical material. In the presentation of the material the main emphasis has been laid upon practical aspects, and some purely mathematical facts are given without proof. On the other hand, in some cases detailed proofs of theorems are given, especially when this elucidates the meaning of the theorem and shows in which way it can be applied. Besides, the study of the proofs helps the student to acquire practice in logical argument and provides prerequisites for further mathematical self-education.

The original variant of this course was inaugurated by the late professor A. F. Bermant, a talented scientist and teacher (died in 1959). Since then all the additions and improvements have been introduced into the course by me.

I am grateful to L. Ya. Tsirlis for his constructive criticism and to professor L. A. Tumarkin for many helpful comments. Their valuable advice has been followed in the preparation of the present edition of the book.

I. G. Aramanovich

CONTENTS

*(The starred items indicate the sections that may
be omitted in a first reading of the book)*

Preface	5
Introduction	15
1. The Subject of Mathematical Analysis	15
2. Variables and Functions	16
3. The Role of Mathematics and Mathematical Analysis in Natural Sciences and Engineering	17
 CHAPTER I. FUNCTION	 19
§ 1. Real Numbers	19
4. Real Numbers and Number Scale. Interval	19
5. Absolute Value	22
6. Approximate Calculations	23
§ 2. The Concept of Function	26
7. Definition of Function	26
8. Methods of Representing Functions	28
9. Notation	31
10. Basic Elementary Functions. Composite Function	32
11. Elementary Functions	33
12. Implicit Functions. Multiple-valued Functions	36
§ 3. Characteristics of Behaviour of Functions. Some Important Examples	38
13. Simplest Characteristics of Functions	38
14. Investigating Function with the Aid of Its Graph	42
15. Direct Variation and Linear Function. Increment	44
16. Quadratic Function	46
17. Inverse Variation. Linear-fractional Function	48
§ 4. Inverse Function. Power, Exponential and Logarithmic Functions	51
18. Inverse Function	51
19. Power Function	55
20. Exponential and Logarithmic Functions	57
§ 5. Trigonometric, Inverse Trigonometric, Hyperbolic and Inverse Hyperbolic Functions	60
21. Trigonometric Functions. Harmonic Vibrations	60
22. Inverse Trigonometric Functions	64
23. Hyperbolic and Inverse Hyperbolic Functions	67
Questions	70
 CHAPTER II. LIMIT. CONTINUITY	 72
§ 1. Limit. Infinitely Large Magnitudes	72
24. Limit of a Function	72
25. Limits of Functions for $x \rightarrow \pm \infty$	75

26. Limit of a Sequence	78
27. Infinitely Large Magnitudes. Bounded Functions	80
28. Infinitesimals	84
29. General Rules for Finding Limits	85
30. Test for the Existence of the Limit of a Function. The Limit of the Function $\frac{\sin x}{x}$ as $x \rightarrow 0$	93
31. Test for the Existence of the Limit of a Sequence. The Limit of the Sequence $\left(1 + \frac{1}{n}\right)^n$ for $n \rightarrow \infty$	95
§ 2. Continuous Functions	98
32. Continuity	98
33. Points of Discontinuity of a Function	100
34. Operations on Continuous Functions. Continuity of Elementary Functions	103
35. Properties of Continuous Functions	107
§ 3. Comparison of Infinitesimals. Comparison of Infinitely Large Magnitudes	109
36. Comparison of Infinitesimals. Equivalent Infinitesimals	109
37. Comparison of Infinitely Large Magnitudes	114
Questions	116
 CHAPTER III. DERIVATIVE AND DIFFERENTIAL. DIFFERENTIAL CALCULUS	
§ 1. Derivative	118
38. Some Physical Problems	118
39. The Rate of Change of a Function at a Point. The Derivative of a Function. The Derivative of the Power Function	121
40. Geometrical Meaning of the Derivative	125
§ 2. Differentiating Functions	126
41. Differentiation and Arithmetical Operations	126
42. Differentiating Composite and Inverse Functions	131
43. Derivatives of Basic Elementary Functions	134
44. Differentiating Elementary Functions. Examples	139
45. Further Remarks on Differentiation of Functions	140
46. Parametric Representation of Functions. Parametric Differentiation	143
§ 3. Some Geometrical Problems. Graphical Differentiation	148
47. Tangent and Normal to a Line	148
48. Graphical Differentiation	151
49. Geometrical Interpretation of the Derivative in Polar Coordi- nates	153
§ 4. Differential	155
50. Differential and Its Geometrical Meaning	155
51. Properties of the Differential	157
52. Differentiable Functions	160
53. Applying Differentials to Approximate Calculations	163
§ 5. Derivatives and Differentials of Higher Orders	167
54. Derivatives of Higher Orders	167
55. Differentials of Higher Orders	171
Questions	173

CHAPTER IV. APPLICATION OF DIFFERENTIAL CALCULUS TO INVESTIGATION OF BEHAVIOUR OF FUNCTIONS	175
§ 1. Theorems of Fermat, Rolle, Lagrange and Cauchy	175
56. Theorems of Fermat and Rolle	175
57. Lagrange's Theorem	178
58*. Cauchy's Theorem	180
§ 2. Investigating Functions with the Aid of First and Second Derivatives	181
59. Testing Functions for Monotonicity	181
60. Extrema of Functions	184
61. Investigating Functions with the Aid of Second Derivatives. The Greatest and the Least Values of a Function	194
62. Convexity and Concavity of a Curve. Points of Inflection	193
§ 3. L'Hospital's Rule. General Scheme for Investigating Functions	204
63. L'Hospital's Rule	204
64. Asymptotes of Curves	211
65. General Scheme for the Investigation of the Graph of a Function	216
§ 4. Curvature	219
66. Differential of Arc Length	219
67. Curvature	221
§ 5. Space Curves. Vector Function of a Scalar Argument	225
68. Space Curves	225
69. Circular Helix	228
70. Vector Function of a Scalar Argument	230
71*. Applications to Mechanics	235
§ 6. Complex Functions of a Real Argument	237
72. Complex Numbers	237
73. Definition of a Complex Function. Differentiation	240
74. Exponential Function. Euler's Formulas	242
§ 7. Solution of Equations	245
75. General Properties of Equations	245
76. Testing a Root for Multiplicity	248
77. Approximate Solution of Equations	249
Questions	255
CHAPTER V. INTEGRAL CALCULUS	258
§ 1. Indefinite Integral	258
78. Antiderivative	258
79. Indefinite Integral. Basic Table of Integrals	261
80. Simplest Integration Rules	263
81. Integration by Parts and by Change of Variable	269
82. Integration of Rational Functions	275
83. Integrating Some Simple Irrational Functions	282
84. Integrating Trigonometric Functions	284
85. Conclusion. Using Tables of Integrals	288
§ 2. Definite Integral	291
86. Some Problems of Geometry and Physics	291
87. Definite Integral. Existence Theorem	295
88. Simplest Properties of the Definite Integral	300
89. Interchanging Limits in the Definite Integral and Splitting the Interval of Integration. Geometrical Meaning of the Integral	302
90. Estimating Definite Integral. Mean Value Theorem. The Arithmetic Mean of a Function	305
91. The Derivative of an Integral with Respect to Its Upper Limit	312
92. The Newton-Leibniz Theorem	314
93*. Integrating Complex Functions of a Real Argument	316

§ 3. Methods of Evaluating Definite Integrals	318
94. Integration by Parts and by Change of Variable in the Definite Integral	318
95. Approximate Integration	323
96. Graphical Integration	329
§ 4. Improper Integrals	331
97. Infinite Integrals	331
98*. Tests for Convergence of Improper Integrals with Infinite Limits	335
99. Integrals of Discontinuous Functions	340
Questions	343
 CHAPTER VI. APPLICATION OF INTEGRAL CALCULUS	 345
§ 1. Some Problems of Geometry and Statics	345
100. Area of a Plane Figure	345
101. Volume of a Solid	348
102. Arc Length	350
103. The Centre of Gravity of a Curvilinear Trapezoid	354
§ 2. General Scheme of the Application of the Integral	358
104. The Problem-solving Scheme	358
105*. The Area of a Surface of Revolution	362
106. Fluid Pressure on the Wall of the Vessel	363
Questions	364
 CHAPTER VII. FUNCTIONS OF SEVERAL VARIABLES AND THEIR DIFFERENTIATION	 365
§ 1. Functions of Several Variables	365
107. Functions of Two Variables. Method of Parallel Sections. Domains and Neighbourhoods	365
108. Limit of a Function of Two Variables. Continuity	372
109. Functions of Several Independent Variables	375
§ 2. Derivatives and Differentials. Differential Calculus	376
110. Partial Derivatives and Differentials	376
111. Total Differential	381
112. Geometrical Interpretation of the Total Differential of a Function of Two Variables	386
113. Applying Total Differential to Approximate Calculations	388
114. Derivatives and Differentials of Higher Orders	390
115. Reconstructing a Function from Its Total Differential	393
116. Differentiating Composite Functions. Rules for Finding Differentials of Functions	400
117. Existence Theorem for an Implicit Function	405
118. Differentiating Implicit Functions	407
§ 3. Applications of Differential Calculus to Geometry	409
119. Surfaces	419
120. Space Curve as Line of Intersection of Two Surfaces	412
§ 4. Extrema of Functions of Two Variables	414
121. Necessary Conditions for Extremum	414
122. Sufficient Conditions for Extremum of a Function of Two Variables	417
123. Determining the Greatest and the Least Values of Functions	419
124*. Conditional Extremum	420
§ 5. Scalar Field	427
125. The Notion of a Scalar Field. Level Surfaces	427
126. Directional Derivative	428
127. Gradient	431
Questions	435

CHAPTER VIII. DOUBLE AND TRIPLE INTEGRALS	437
§ 1. Double Integrals	437
128. The Volume of a Curvilinear Cylinder. The Double Integral	437
129. Properties of Double Integrals	440
130. Evaluating Double Integrals	442
131. Change of Variables in the Double Integral. Polar Coordinates	450
132. Applications of Double Integrals to Problems of Mechanics	457
§ 2. Triple Integrals	459
133. The Mass of a Nonhomogeneous Body. The Triple Integral	459
134. Evaluating Triple Integrals	461
135. Applications of Triple Integrals	468
§ 3*. Integrals Dependent on Parameters	471
136*. Integrals with Finite Limits	471
137*. Improper Integrals Dependent on Parameters. The Gamma Function	476
Questions	481

CHAPTER IX. LINE INTEGRALS AND SURFACE INTEGRALS. FIELD THEORY	482
§ 1. Line Integrals	482
138. The Work of a Field of Force. The Line Integral	482
139. Evaluating Line Integrals. Integral over a Closed Contour	485
140. Green's Formula	491
141. Conditions for a Line Integral Being Path-independent	494
142. Integrating Total Differentials. Primitive	501
143. Line Integrals over Space Curves	503
144. Application of Line Integrals to Problems of Mechanics and Thermodynamics	507
145. Line Integrals with Respect to Arc Length	512
§ 2*. Surface Integrals	515
146*. Fluid Flux Across a Surface. The Surface Integral	515
147*. Properties of Surface Integrals	519
148*. Evaluating Surface Integrals	522
149*. Stokes' Theorem	528
150*. Gauss-Ostrogradsky Formula	531
§ 3*. Field Theory	533
151*. Vector Field and Vector Lines	533
152*. Flux of a Vector Field. Divergence	536
153*. Circulation and Rotation of a Vector Field	542
154*. Hamiltonian Operator. Repeated Application of the Hamiltonian Operator to a Vector Field	547
155*. Properties of Some Vector Fields	550
156*. Electromagnetic Field	554
157*. Nonstationary Fields	559
Questions	561

CHAPTER X. DIFFERENTIAL EQUATIONS	564
§ 1. Differential Equations of the First Order	564
158. General Notions. Existence Theorem	564
159. Equations with Variables Separable	568
160. Some Physical Problems	571
161. Homogeneous Equations of the First Order. Linear Equations	574
162. Exact Differential Equations	581

163. Approximate Methods of Solving First-Order Differential Equations	582
164*. Singular Points and Singular Solutions. Envelope	585
§ 2. Differential Equations of the Second and Higher Orders	592
165. Differential Equations of the Second Order	592
166. Some Particular Types of Equations of the Second Order	595
167. Applications to Mechanics	597
168. Higher-order Differential Equations	602
§ 3. Linear Differential Equations	603
169. Linear Equations of the Second Order. General Properties	603
170. Homogeneous Linear Differential Equations of the Second Order with Constant Coefficients	607
171. Nonhomogeneous Linear Differential Equations of the Second Order with Constant Coefficients. The Method of Complex Amplitudes	612
172. The Method of Variation of Arbitrary Constants	619
173. Linear Differential Equations of the n th Order	622
174. Linear Differential Equations of the n th Order with Constant Coefficients	625
175. Vibrations. Resonance	627
§ 4. Systems of Differential Equations	635
176. General Definitions. Normal Form of a System of Differential Equations	635
177*. Geometrical and Mechanical Interpretation of Solutions of a System of Differential Equations. Phase Space	639
178. Systems of Linear Differential Equations	642
179. Systems of Linear Differential Equations with Constant Coefficients	644
180*. Multiple Roots of the Characteristic Equation	649
181*. Matrix Form of a System of Linear Differential Equations	652
Questions	656
 CHAPTER XI. SERIES	 659
§ 1. Numerical Series	659
182. Series. Definition of the Sum of a Series	659
183. Necessary Condition for Convergence of a Series. Harmonic Series	664
184. Positive Series. Sufficient Conditions for Convergence	665
185. Cauchy's Integral Test	670
186. Series with Arbitrary Terms. Absolute Convergence	674
§ 2. Functional Series	678
187. General Definitions	678
188. Properties of Regularly Convergent Functional Series	681
§ 3. Power Series	684
189. Abel's Theorem. Interval of Convergence and Radius of Convergence	684
190. Properties of Power Series	689
§ 4. Expanding Functions into Power Series	691
191. Taylor's Series	691
192. Key Condition for Expanding a Function into Taylor's Series	694
193. Taylor's Formula. Lagrange's Form of the Remainder in Taylor's Formula	695
194. Expanding Functions into Taylor's and Maclaurin's Series	699
§ 5. Some Applications of Taylor's Series	706
195. Approximate Computation of Values of Functions	706

196. Integrating Functions. Application of Taylor's Series to Differential Equations	711
§ 6*. Some Further Topics in the Theory of Power Series	716
197*. Power Series in Complex Argument	716
198*. Taylor's Series and Taylor's Formula for Functions of Two Independent Variables	719
Questions	721
CHAPTER XII. FOURIER SERIES AND FOURIER INTEGRAL . . .	724
§ 1. Fourier Series	724
199. Harmonic Vibrations. Trigonometric Series	724
200. Fourier Series	728
201. Expanding Even and Odd Functions into Fourier Series. Fourier Series for an Arbitrary Interval	732
202*. Fourier Series in Complex Form	740
203. Practical Harmonic Analysis	742
§ 2. Some Further Topics in the Theory of Fourier Series	746
204*. Orthogonal Systems of Functions	746
205. Mean Square Deviation. Minimizing Property of Fourier Coefficients	748
206. Convergence in the Mean. Parseval's Relation. Lyapunov's Theorem	751
§ 3*. Fourier Integral	753
207*. Fourier Integral	753
208*. Fourier Integral for Even and Odd Functions	756
209*. Fourier Integral in Complex Form. Fourier Transformation	758
Questions	761
Table of Integrals	762
Bibliography	768
Name Index	770
Subject Index	772

Introduction

1. **The Subject of Mathematical Analysis.** Mathematical analysis is an important division of higher mathematics. It should be stressed at the very beginning that the classification of mathematics into "elementary" and "higher" is rather conventional and has no exact criteria.

The divisions of mathematics which are traditionally included into elementary mathematics appeared very long ago. For instance, the discovery of the basic facts studied at schools in the elementary course of geometry is connected with the names of the Greek scientists *Pythagoras* (5th century B.C.) and *Euclid* (3rd century B.C.). Algebra as an independent science appeared in the 8th century A.D.; its foundations were presented by the Uzbek scientist *Mohammed Al-Khorezmi* in his famous treatise. Of course, the basic rules of arithmetical and algebraic operations had been known much earlier. As to trigonometry whose development was closely connected with astronomy, its fundamentals were known in antiquity. But since then algebra and trigonometry underwent significant changes and assumed their modern form comparatively not long ago. For instance, the systematic usage of algebraic notation and symbols is due to the book of the French mathematician *F. Vieta* published in 1591.

The fundamentals of those divisions of mathematics which are attributed to higher mathematics were created in the 17th and 18th centuries in connection with the general progress of natural science and industry. The needs of science and industry in a deeper understanding and exploration of phenomena of nature led to the investigation of *processes* and *motions* in the real world. This development was first of all connected with studying physical phenomena since their quantitative investigation required mathematical methods capable of describing interrelations between the variations of the quantities involved.

Mathematical analysis is one of the most important divisions of higher mathematics; its main object is studying *variables and their relationships*. It is closely related to and based on the connection between algebraic and geometrical methods which for the

first time appeared in *analytic geometry* created by the French mathematician and philosopher *R. Descartes* (1596-1650). The founders of the *differential and integral calculus* forming the foundation of mathematical analysis were *I. Newton* (1642-1727), the great English scientist, and *G. Leibniz* (1646-1717), the great German philosopher and mathematician.

2. Variables and Functions. The fundamental idea of the application of mathematical methods to natural sciences and engineering is connected with the basic concept of a *magnitude*. *We regard as a scalar magnitude everything that can be measured and expressed by a number.*

In concrete problems of natural and technical sciences we encounter magnitudes of various types. Examples of *scalar* magnitudes are length, area, volume, mass, temperature, etc. Besides scalar magnitudes we also deal with *vector* quantities such as force, velocity, acceleration and the like. Mathematical analysis begins with studying operations on scalar magnitudes which are simpler than vectors.

When stating mathematical propositions and laws we usually abstract from the concrete physical nature of the magnitudes involved and take into account only their *numerical values*. Accordingly, *mathematics deals with a general notion of a magnitude without considering its physical meaning*. This fundamental feature of mathematics, its *abstractness*, is extremely important since it makes it possible to apply mathematical methods to various kinds of activity and phenomena and provides its generality and universality. The abstractness of mathematics is a powerful tool in practical work and has nothing in common with the indifference to reality.

Usually, in a group of magnitudes involved in a problem there are some which change while others remain invariable. *Variation* is one of the most important features characterizing what we call a *motion* or a *process*. A natural phenomenon or an industrial process is usually observed as a *variation of some magnitudes involved which is produced by the change of the others*. For instance, if the volume of a mass of gas kept at a constant temperature varies, its pressure also undergoes a variation. Similarly, in a free fall of a body (in vacuum) under the action of gravity we observe the variation of the velocity of motion as the distance between the body and its initial position changes and also the variation of this distance and of the velocity of motion in time. At the same time, the acceleration of gravity remains constant at any time moment along the whole path.

Therefore, to apply mathematics to studying natural phenomena and other kinds of processes it is necessary to introduce the notion of a *variable*.

If a magnitude assumes different numerical values we call it a variable. If a magnitude retains one and the same value we call it a constant.

As has been mentioned, from the quantitative point of view, every process is characterized by a mutual variation of a number of magnitudes. This leads to the most important mathematical concept of a *functional relationship*, i.e. to the idea of an interconnection between variables.

The main purpose of a natural or technical science is to establish the relationships between the variables involved in the process under consideration and to describe it mathematically. *The law of a process is nothing but a functional relationship observed in this process and characterizing it.* For instance, the functional relationship between the pressure (p) and the volume (v) of a mass of gas having a constant temperature is described by the equality

$p = \frac{k}{v}$ where k is a constant and expresses the general law which the gases obey in the corresponding circumstances. Its verbal formulation saying that the *gas pressure at a constant temperature is inversely proportional to the volume* is an ordinary statement of that law. Similarly, the functional relationship between the path length (s) covered by a freely falling body in vacuum and the time taken (t) is described by the formula $s = \frac{1}{2}gt^2$ where g is the acceleration of gravity and expresses the general *law of free fall*.

The basic and the most important aim of mathematical analysis is to provide a thorough investigation of functional relationships.

3. The Role of Mathematics and Mathematical Analysis in Natural Sciences and Engineering. In mathematics and, in particular, in mathematical analysis practical work and observation of nature are, as in other sciences, the main source of scientific discoveries. In their turn, mathematical methods play a very important role in natural sciences and engineering. Mathematical methods lie in the foundation of physics, mechanics, engineering and other natural sciences. For all of them mathematics is a powerful theoretical and practical tool without which no scientific calculation and no engineering and technology are possible. Mathematical analysis which treats of variables and functional relationships between them is particularly important since the laws of physics, mechanics, chemistry, etc. are expressed as such relationships.

An important feature of the application of mathematics to other sciences is that it enables us to make scientific *predictions*, that is to draw, on the basis of logic and with the aid of mathematical methods, correct conclusions whose agreement with reality is then confirmed by experience, experiment and practice. Here

we confine ourselves to two remarkable examples illustrating what has been said.

As is known, the modern science of aviation was created by the famous Russian scientist Professor *N.E. Zhukovsky* (1847-1921). He derived by means of mathematical methods certain formulas and laws which enabled him to predict the possibility of aerobatics and, in particular, of looping the loop. Soon the loop was in fact performed by the Russian pilot, captain *P.N. Nesterov*. The possibility of looping the loop was thus discovered mathematically before it was realized physically!

Here is another example. *J. Leverrier*, a French astronomer, studied planetary motion in the solar system. When applying the laws of classical mechanics expressed in the form of some known functional relationships he discovered a discrepancy between the theoretical results and real observation. Then he found that the discrepancy can be eliminated if the existence of an unknown planet possessing a certain mass and moving in a certain orbit is assumed. Soon this new planet (Neptune) was in fact found (in 1846) exactly at the time moment and position predicted by Leverrier who thus discovered a new heavenly body by means of calculations made on a piece of paper at his desk! Nowadays astronomical predictions are taken for granted but they are only possible because the techniques of mathematics and mathematical analysis are based on our knowledge of objective laws of reality.

In recent years the role of mathematics has still increased especially in connection with the appearance of modern high-speed electronic computers. Realization of space flights, launching rockets to other planets and establishing radio and television communication with them require extremely complicated and precise mathematical calculations which cannot be performed without computers. Mathematical methods are penetrating deeply even into such traditionally "non-mathematical" sciences as economics, biology, medicine, etc. It can be asserted without exaggeration that no modern scientific and technical project can be realized without resorting to mathematics and its methods.

Chapter I

FUNCTION

§ 1. Real Numbers

4. Real Numbers and Number Scale. Interval. We begin the course of mathematical analysis with *real numbers*. Let us recall some basic definitions.

The positive whole numbers 1, 2, 3, ... are called *natural*.

Adding to the natural numbers all the positive fractions and zero and also all the negative integers and fractions we obtain the set¹ of all *rational numbers*. Every rational number has the form $\frac{p}{q}$ where p and q are integers and $q \neq 0$.

Every rational number can be written either as a finite decimal or as an infinite decimal with an indefinitely repeating finite block of digits, both types of decimals being referred to as *periodic decimals*.

The infinite decimals which are not repeating (nonperiodic decimals) represent *irrational numbers*. It can be proved, for instance, that the numbers $\sqrt{2}$, $\sqrt{3}$, $\log 3$, π , $\sin 20^\circ$ are irrational.

All rational and irrational numbers form the set of real numbers.

Now we proceed to the geometric representation of numbers. Let us take a straight line and a point O on it as the initial point (the origin) for reckoning lengths. We also choose a scale, that is unit length, and a positive direction along the line.

Definition. A straight line with origin, scale and positive direction for reckoning lengths is called a *number scale* (or a *number line*).

We usually think of a number scale as a horizontal straight line with positive direction from left to right.

Let us lay off from the origin all the line segments *commensurable* with unit length; as is known from geometry, the lengths

¹ The notion of "set" is considered as primitive and well-understood and is not defined. Examples of sets as collections of certain elements are given in the text of the book. Most often we deal with sets of numbers or points.

of these segments are expressed by rational numbers. According to the chosen positive direction, the line segments set off to the right of the point O are regarded as corresponding to positive numbers and those set off to the left of O as corresponding to negative numbers.

Suppose that the end point of a line segment which, relative to the chosen scale, corresponds to a rational number a falls at a point M of the line. We then say that the point M represents the number a or, synonymously, that the number a is the *coordinate* of the point M . Then obviously the point O represents the number zero.

The points of the number scale representing the rational numbers are called *rational points*.

Imagine that we consider the points representing the rational numbers of the form $\pm \frac{1}{N}, \pm \frac{2}{N}, \dots$ with indefinitely growing N ; then the "density" of the set of representing points on the number scale also increases. Moreover, we can show that, given any two different rational points A and B located arbitrarily close to each other on the number scale, there exists an infinitude of rational points lying between them. Indeed, the points A and B have rational coordinates a and b . The midpoint (denote it by C) of the line segment with the end points a and b has the coordinate $\frac{a+b}{2}$ which is also a *rational* number. The same argument implies that the midpoint C_1 of the line segment AC and the midpoint C_2 of the segment CB are also rational points. Since we can proceed in this way indefinitely, it becomes clear that there are infinitely many rational points on the line segment AB .

And yet the rational points do not exhaust the number scale; this means that there are also irrational points on it. For, if we lay off from the point O a line segment whose length is equal to that of the diagonal of a square with unit side, the end point of the segment falls at a point which is not rational since, as is known, the diagonal of unit square is *incommensurable* with unit length and cannot be expressed by any rational number. Generally, the ends of all segments set off from the origin which are *incommensurable* with the scale (their lengths are expressed by irrational numbers) lie at points which are not rational. We shall call these points *irrational*.

All the rational and irrational points no longer "leave gaps" on the number scale, that is, they completely exhaust it. It should be stressed that there is a *one-to-one correspondence* between the set of all points of the number scale and the set of all real numbers: *every point of the number scale represents one real num-*

ber (which can be rational or irrational) and, conversely, every real (rational or irrational) number is a coordinate of exactly one point of the number scale.

The number scale is convenient for representing, in a visual manner, various relations between numbers. For instance, if one number is smaller than another the point representing the former lies to the left of the point representing the latter, and to a number contained between two given numbers there corresponds a point lying between the representing points of those two numbers.

The one-to-one correspondence between the real numbers and the points of the number scale makes it possible to identify them in the sense that we can make no distinction between a number and a point and interpret, depending on the situation, a point as a number and a number as a point.

Definition. The set of all numbers (points) contained between two given numbers (points) is called an *interval*, these two numbers (points) being referred to as the *end points* of the interval.

An interval with end points $x=a$ and $x=b$ where $a < b$ can be specified by the inequalities $a < x < b$; its commonly used notation is (a, b) . If the end points of an interval (a, b) are added to the set of all its points the resultant set of points (numbers) is spoken of as a *closed interval*. The closed interval with end points $x=a$ and $x=b$ is determined by the inequalities $a \leq x \leq b$ and designated by the symbol $[a, b]$. In contrast to the closed interval $[a, b]$, the interval (a, b) is called *open* (or non-closed). If one of the end points is included into the interval while the other is not, the resultant set is specified by the inequalities $a \leq x < b$ (if the end point a is added to the interval) or by $a < x \leq b$ (if the end point b is joint to the interval). Respectively, the "semiclosed" intervals thus obtained are denoted as $[a, b)$ and $(a, b]$.

When there is no need in distinguishing whether or not an end point is included into the interval in question we simply speak of an *interval*.

Besides finite intervals discussed above we often deal with *infinite intervals*. An (open) infinite interval is the set of all numbers less than a given number c or the set of numbers greater than c , or the set of all real numbers. These infinite intervals are, respectively, specified by the inequalities $-\infty < x < c$, $c < x < +\infty$ and $-\infty < x < +\infty$. The notation is, respectively, $(-\infty, c)$, $(c, +\infty)$ and $(-\infty, +\infty)$.

If in the first two cases the point c is included into the interval the resultant sets are denoted as $(-\infty, c]$ and $[c, +\infty)$.

Definition. An open interval of length $2l$ with centre at a point a is called an *l -neighbourhood* of the point a

The coordinates x of the points belonging to the l -neighbourhood of the point a (see Fig. 1) obviously satisfy the inequalities

$$a-l < x < a+l$$

The notion of a neighbourhood of a point will be frequently used in our course.

If a point x belongs to a given interval this is designated with the aid of the *inclusion relation* $\in$. For instance, the notation

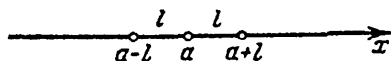


Fig. 1

$x \in [a, b]$ means that x is one of the points of the closed interval $[a, b]$, or, which is the same, the number x satisfies the inequalities $a \leq x \leq b$. Similarly, $x \in (a, b)$ means that $a < x < b$, etc.

An analogous inclusion relation $\subset$ is used for expressing the fact that one set is a part (a *subset*) of another set; for instance, $[a, b] \subset (c, d)$ means that all the points of the closed interval $[a, b]$ belong to the open interval (c, d) , etc.

5. Absolute Value.

Definition. The *absolute value* (or the *modulus*) of a number A is a number, denoted as $|A|$, which is equal to A , if A is positive or equal to zero, and equal to $-A$ if A is negative:

$$|A| = \begin{cases} A & \text{if } A \geq 0 \\ -A & \text{if } A < 0 \end{cases}$$

Hence, $|A|$ is always a nonnegative number (that is positive or zero). $|A|$ is equal to the length of the line segment connecting the origin with the point representing the number A on the number scale.

The relation

$$|A| \leq B \text{ where } B > 0 \quad (*)$$

is equivalent to the two relations

$$-B \leq A \leq B \quad (**)$$

For inequality $(*)$ means that the point A is at a distance less than B from the origin, which is only possible when this point belongs to the interval $(-B, B)$, i.e. belongs to the B -neighbourhood of the origin, which implies inequalities $(**)$. The converse is also obvious.

The condition that x belongs to an l -neighbourhood of a point a (which is expressed by the inequalities $a-l < x < a+l$) can now be written as $|x-a| < l$.

Let us proceed to establish the following propositions on absolute values:

(1) *The absolute value of a sum does not exceed the sum of the absolute values of the summands, that is*

$$|A+B+C+\dots+D| \leq |A|+|B|+|C|+\dots+|D| \quad (***)$$

Let us begin the proof with the case of two summands. It can be checked directly that if A and B are of the same sign the left-hand and the right-hand members of relation (***) are equal and that if their signs are different the former is less than the latter. Now, passing, in succession, from two summands to three, from three to four and so on we readily prove the desired general assertion. The sign of equality only appears in relation (***) when all the summands are of the same sign.

(2) *The absolute value of a difference is not less than the difference of the absolute values of the minuend and the subtrahend:*

$$|A-B| \geq |A|-|B|$$

To prove the assertion let us put $A-B=C$. Then $A=B+C$ and, according to the above,

$$|A|=|B+C| \leq |B|+|C|=|B|+|A-B|$$

that is

$$|A-B| \geq |A|-|B|$$

(3) *The absolute value of a product is equal to the product of the absolute values of the cofactors:*

$$|A \cdot B \cdot C \dots D| = |A| \cdot |B| \cdot |C| \dots |D|$$

(4) *The absolute value of a quotient is equal to the ratio of the absolute values of the dividend and the divisor:*

$$\left| \frac{A}{B} \right| = \frac{|A|}{|B|} \text{ if } B \neq 0$$

The last two assertions are quite evident.

6. *Approximate Calculations.* In engineering sciences and practical work we almost always deal with numbers expressing *approximate* numerical values of the magnitudes involved.

We know that a number of successive identical measurements of one and the same magnitude always result in a sequence of different values (although the discrepancy may be very small). When the measure of a concrete magnitude is said to be equal to a given number it is simply meant that the corresponding measurements give results for which the discrepancy is so small that it can be neglected.

In the applications of mathematical analysis to natural and engineering sciences we always have to carry out certain calcula-

tions, and the final result is expressed numerically and obtained by means of calculations. Therefore it is absolutely necessary to be able to perform accurate calculations. But in what practical sense should the term "accurate calculations" be understood?

First of all the result must be of course sufficiently *precise*. But, on the other hand, another condition which is, to a certain extent, opposite to the above, should hold: the accuracy must not be *excessive* or *fictitious*. And, finally, it is desirable to carry out the calculations *speedily* with minimum labour and time. When calculating one should always keep in mind these requirements.

Now we proceed to define basic notions related to approximate calculations.

Let A denote the precise value of a magnitude which is being measured or, simply, a number.

Definition. A number a is called an *approximate* value of A with an error α if

$$A - a = \alpha$$

For an approximate equality we shall use the symbol $\approx$:

$$A \approx a$$

The number a can be less than A (i.e. $a < A$) or greater than A ($a > A$). In the former case $\alpha > 0$, and a is a *minor approximation* of A , and in the latter case $\alpha < 0$, and a is a *major approximation* of A . It does not often matter which of these two cases takes place.

The error α and the precise value of A are usually unknown, and therefore it is important to determine a *bound* ε which cannot be exceeded by the absolute value of the error: $|\alpha|$ is not greater than ε , that is $|\alpha| \leq \varepsilon$.

Definition. The absolute value $|\alpha|$ of the error is called the *absolute error*. If it is known that the absolute error $|\alpha|$ of an approximate value a does not exceed a positive number ε , this number ε is termed the *limiting absolute error* of the approximate value a :

$$|A - a| \leq \varepsilon$$

or, equivalently,

$$-\varepsilon \leq A - a \leq \varepsilon$$

The number ε is not determined uniquely since every number greater than a given absolute error can serve as a limiting absolute error and can be taken as the number ε .

For instance, the number 3.24 is an approximation to the number 3.2447 having an absolute error of 0.0047, and at the same time it is an approximate value of 3.2447... with a limiting

absolute error of 0.005 since $0.0047... < 0.005$. By the way, even when the error is known it is sometimes more convenient to indicate the limiting error; for instance, in this example it is more convenient to indicate the limiting error 0.005 instead of the error 0.0047.

Suppose that we carry out a measurement with the aid of a simple measuring instrument. Then it is often possible to determine the limiting absolute error. If, for instance, a length is measured with a ruler with graduation mark (scale division) of one centimetre it is clear that the error does not exceed 0.5 cm. Here of course we only mean the error introduced when reading the data from the ruler. It is generally assumed that the limiting absolute error of reading the meter of a measuring instrument does not exceed half the graduation mark.

For brevity, the word "limiting" is often omitted and the limiting absolute error ϵ is simply called the absolute error.

The smaller ϵ , the more accurate the approximate equality $A \approx a$.

It should be noted that it is impossible to judge upon the accuracy of an approximation by the absolute error itself; to estimate the accuracy we must have at least approximate information about the value of A . For instance, an absolute error of 0.5 m indicates an imperfect measurement of the height of a room while it indicates a sufficient accuracy in a measurement of the dimensions of a lot of land; if the same absolute error is obtained when measuring the altitude of a mountain the accuracy of the measurement should be considered very high. The quality of an approximation is usually judged upon by the ratio of the error to the value of the magnitude in question, that is by the *relative error* understood as the absolute value of the ratio of the error α to the exact value of A .

Since usually we only know a *limiting absolute error* ϵ and an *approximate value* a of the magnitude A the relative error is estimated with the help of the *limiting relative error*.

Definition. The *limiting relative error* δ is the ratio of the limiting absolute error to the approximate value a of the magnitude A :

$$\delta = \frac{\epsilon}{|a|}$$

As above, the word "limiting" is usually omitted and the term "relative error" is understood as the limiting relative error. The smaller the limiting relative error, the better the approximation and the more accurate the measurement or calculation.

Relative errors are usually expressed in per cent. To obtain the percentage from a given relative error the latter should be multiplied by 100.

(2) one or several intervals (finite or infinite) of the number scale.

In the former case we speak of a *function of an integral argument* and in the latter of a *function of a continuous argument*.

The set of values assumed by a function y is called the *range* of the function.

Let us consider some examples of functions.

1. The *air temperature* τ at a given place measured during the day is a function of *time* t since to every value of the variable t there corresponds a definite value of the magnitude τ .

2. The *perimeter* P of a regular polygon inscribed in a given circle is a function of an integral argument which is the *number* n of sides of that polygon. The domain of definition of this function is the set of all positive integers n greater than or equal to 3. Another example of a function of the same argument is the *area* S of that regular polygon.

3. The *area* S of a circle is a function of its *radius* r which can take on any positive value. To every value of r belonging to the interval $(0, \infty)$ there corresponds a definite value of S determined by the formula

$$S = \pi r^2$$

The domain of definition of this function is the set of all positive real numbers, that is the positive semi-axis $0 < r < \infty$.

4. The *distance* s travelled by a body freely falling in vacuum is a function of the *time* t of the fall. To every value of t there corresponds a definite value of s specified by the formula

$$s = \frac{1}{2} g t^2$$

where g is the acceleration of gravity. The domain of definition of this function is the interval $[0, T]$ where T is the time of fall.

We see that the dependence of the distance travelled in a free fall of a body on the time taken is of the same character as the dependence of the area of a circle on its radius: in both cases the function is *directly proportional to the square of the independent variable*. From the point of view of mathematics this is the only important feature of these relationships since in mathematics we are interested in the way in which the function depends on its argument but not in the concrete physical phenomenon described by the given functional relationship. In mathematical analysis the functions discussed in Examples 3 and 4 are simply regarded as particular cases of the so-called *quadratic function* (see Sec. 16).

The definition of a function does not require that the function should necessarily vary as the independent variable changes. For instance, in Example 1 it may turn out that the air temperature τ remains constant during a time period t , but this does not prevent us from considering the temperature as a function of time.

Let us consider another example. The function

$$y = a \cos^2 x + b \sin^2 x$$

where a and b are constants, takes different values depending on the values of x if $a \neq b$. But if $a = b$ the function y is a constant equal to a for all the values of x .

Thus we can say, including the latter particular case, that the expression $a \cos^2 x + b \sin^2 x$ is a function of x for *any* values of the constants a and b .

8. Methods of Representing Functions. To represent or, which is the same, to specify a function means to indicate its domain of definition and the rule according to which for any given value of the independent variable the corresponding value of the function is found. The basic methods of representing functions are the *tabular* method, the *analytical* method (the representation by means of a *formula*) and the *graphical* method.

1. Tabular method. When specifying a function by means of a table we simply write down a sequence of values of the independent variable and the values of the function corresponding to them. This way of representing functions is widely used; for instance, the reader is undoubtedly familiar with tables of logarithms, tables of trigonometric functions and their logarithms, etc.

The tabular method is particularly often used in natural sciences and engineering. The numerical results obtained in a sequence of observations of a process are usually compiled in a table. For example, suppose that when studying the dependence of the electrical resistance r of a copper rod on the temperature t° we obtain the following table:

t°	19.1	25.0	30.1	36.0	40.0	45.1	50.0
r	76.30	77.90	79.75	80.80	82.35	83.90	85.10

The resistance is a function of the temperature, and the table specifies the values of this function for the given values of the independent variable.

An advantage of a tabular representation of a function is that for any value of the independent variable included into the table

If, for example, we want to apply the methods of mathematical analysis to investigating the functional relationship between the temperature and electrical resistance of the copper rod, discrete values of the magnitudes involved, even compiled into a table of large capacity, are insufficient for this purpose, and a formula connecting the values of the independent variable and of the function is needed.

The analytical method also possesses some disadvantages:

(a) the lack of visuality;

(b) the necessity of calculations which sometimes may be lengthy.

III. Graphical representation. We begin with the following definition.

Definition. The *graph* of a function (in rectangular Cartesian coordinates) is the locus of all points whose abscissas are the values of the independent variable and the ordinates are the corresponding values of the function.

In other words, if we take the abscissa equal to a value of the independent variable the ordinate of the corresponding point of the graph is equal to the value of the function corresponding to that value of the independent variable. When plotting a graph we can take similar or different scales along the coordinate axes.

Usually the graph of a function is a *curve*¹ (which, in a particular case, can be a straight line). For instance, the graph of the quadratic function $y=x^2$ is a parabola whose axis coincides with the positive half of the axis of ordinates; the graph of the function $y=\frac{1}{x}$ is an equilateral (rectangular) hyperbola for which the coordinate axes are the asymptotes, etc. The graph of a constant function is, apparently, a straight line parallel to the axis of abscissas.

Conversely, if every curve in the plane Oxy such that any straight line parallel to the axis of ordinates has no more than one common point with it serves as the graph of a function: the values of this function are equal to the ordinates of the points of the curve corresponding to the abscissas equal to the values of the independent variable.

Hence, the notion of a *curve* and that of a *function* are closely interrelated. The specification of a function generates a curve which is the graph of that function; conversely, the specification of a plane curve generates a function for which this curve serves as its graph. The graphical representation of a function consists in specifying its graph.

In physics and technical sciences functions are often specified graphically, and in many cases this is the only possible way of

¹ The graph of a function of a discrete argument is a set of disjoint points, in the xy -plane, with integral abscissas.

describing the function. This is most often the case when a self-recording apparatus is used whose tracer automatically records the variation of one magnitude depending on the other (which, for instance, can be time). This results in the appearance of a curve on the tape of the apparatus which is the graph of the functional relationship recorded in the experiment or measurement. For example, a barograph, a device of this type, records a barogram, that is the graph of the atmospheric pressure, and a thermograph traces the graph of the temperature as a function of time.

The graph of a function, like a table, does not provide facilities for a direct application of the methods of mathematical analysis, and this is a demerit of the graphical method; but, at the same time, it has an important advantage: the graphical representation is *visual* and demonstrates directly the character of the variation of the function.

9. Notation. To express the fact that a variable y is a function of a variable x we use the notation

$$y = f(x)$$

(read " y is equal to f of x "). Here the letter f symbolizes the law of correspondence between the independent variable and the function. Besides f any other letter can be used for this purpose: F , φ , Φ , etc. We sometimes also write $y = y(x)$. The symbol of functional relationship is followed by the symbol of the independent variable written in the parentheses.

Some well-known and widely used functions are denoted by certain standard symbols: $\log x$, $\sin x$, $\arccos x$, etc.

As was said, in the relation $y = f(x)$ the symbol f designates the rule of correspondence, and in the case of an analytical representation of a function it also indicates a set of mathematical operations (and their order) which should be performed on x to obtain y . For instance, in the formula

$$f(x) = \sqrt{x^2 + \sin x}$$

the symbol f means that for the given value of x its square and the value of the sine of x are first found after which the results are added together and the square root is extracted from the sum.

If we simultaneously deal with two functions for which the functional relationships between the variables involved are the same, they can be denoted by a common symbol. But if the relationships differ the symbols must also be different. For example, the area S of a circle depends on its radius r like the surface area Q of a sphere on its diameter d , and therefore in both cases we can write

$$S = f(r) \quad \text{and} \quad Q = f(d)$$

using one and the same symbol f . But if we want to express the fact that the circumference $l=2\pi r$ and the area $S=\pi r^2$ of a circle are functions of its radius r we cannot use one and the same symbol and must use two different symbols; for instance, we can write

$$l=f(r) \quad \text{and} \quad S=\varphi(r)$$

If we take a concrete value of the independent variable and substitute it into the expression of the function we obtain a *particular (concrete)* value of the function. Let the particular value of a function $y=f(x)$ corresponding to $x=a$ be equal to b . Then we write

$$b=f(a)$$

The latter relation is no longer understood in the sense that b is a function of a since both a and b are constants. In this case we should say " b is equal to the

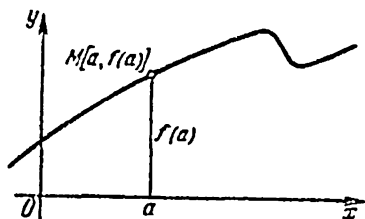


Fig. 2

value of the function f for the value of the independent variable equal to a "; for brevity it is usually said " b is equal to f of a ", but one should bear in mind the precise meaning of the statement indicated here. The notation $y|_{x=a}=b$ is also frequently used in such cases.

The point M of the graph of a function $y=f(x)$ corresponding to the abscissa $x=a$ has the ordinate $f(a)$, which is written as $M[a, f(a)]$ (Fig. 2).

10. Basic Elementary Functions. Composite Function.

Definition. The class of *basic elementary functions* includes the following functions:

- (1) the *power function* $y=x^a$ where the exponent a is real;
- (2) the *exponential function* $y=a^x$ where a is a positive number different from 1;
- (3) the *logarithmic function* $y=\log_a x$ to the base a which is a positive number different from 1;
- (4) the *trigonometric functions* $y=\sin x$, $y=\cos x$, $y=\tan x$ and also the less frequently used functions $y=\sec x$, $y=\csc x$ and $y=\cot x$;
- (5) the *inverse trigonometric functions* $y=\arcsin x$, $y=\arccos x$, $y=\arctan x$ and also $y=\operatorname{arcsec} x$, $y=\operatorname{arccsc} x$ and $y=\operatorname{arccot} x$.

These functions are well known from the elementary course of mathematics studied at schools, and we shall briefly discuss them in § 3.

The basic elementary functions can be used to construct more complicated functions by means of arithmetical operations (ad-

dition, subtraction, multiplication and division) and a new operation of forming "a function of a function" to which we proceed now.

Let a function $u = \varphi(x)$ of the argument x be defined on a set D , G being the range of that function u . Furthermore, let another function $y = f(u)$ be defined on the set G . Then, to each value of x belonging to the set D there corresponds a definite value of u belonging to the set G , and to u , in its turn, there corresponds a definite value of y . Therefore, ultimately, to each value of x from the set D there corresponds a definite value of y , and hence y is a function of x . Denoting this new function by $y = F(x)$ we can write down the expression of the function $F(x)$ in terms of the functions f and φ :

$$y = F(x) = f[\varphi(x)]$$

It is said that $F(x)$ as function of x is a *composite function*¹ ("a function of a function") formed of the functions f and φ . The function $u = \varphi(x)$ entering into the expression $f[\varphi(x)]$ is referred to as the *intermediate variable*.

Thus, the argument of a function can be an independent variable or an intermediate variable dependent on another independent variable. It is the latter case when a composite function appears. For instance, the function $y = \sin^2 x$ can be regarded as a composite function of x since y can be interpreted as a quadratic function of an intermediate variable which, in its turn, is equal to the sine of the independent variable x : if the intermediate variable is denoted by u we can write $y = u^2$ where $u = \sin x$.

A composite function can be obtained by applying the operation of forming a function of a function to more than two constituent functions. A given function can be represented as a composite one in many ways, and an appropriate choice of such a representation often facilitates the application of the techniques of mathematical analysis.

11. Elementary Functions.

I. The functions constructed of *basic* elementary functions and constants by means of a finite number of arithmetical operations and operations of forming a function of a function are called *elementary functions*.

For example, all the functions written below are elementary:

$$y = \frac{a^2}{a + 2b \tan^2 x} + 3c \sin^2 x, \quad y = \frac{2 + \sqrt[3]{x}}{3 - \sqrt{x}}$$

$$y = \log(x + \sqrt{1 + x^2}), \quad y = \frac{10^x - 1}{x^2}, \quad y = \arctan \sqrt{\frac{1 + \sin x}{1 - \sin x}}$$

¹ The term "composite function" characterizes only the way in which the analytical expression of the function is obtained and does not mean that this function necessarily has a complex structure in the ordinary sense.

using one and the same symbol f . But if we want to express the fact that the circumference $l=2\pi r$ and the area $S=\pi r^2$ of a circle are functions of its radius r we cannot use one and the same symbol and must use two different symbols; for instance, we can write

$$l=f(r) \quad \text{and} \quad S=\varphi(r)$$

If we take a concrete value of the independent variable and substitute it into the expression of the function we obtain a *particular (concrete)* value of the function. Let the particular value of a function $y=f(x)$ corresponding to $x=a$ be equal to b . Then we write

$$b=f(a)$$

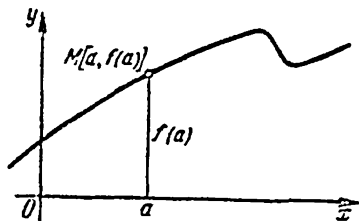


Fig. 2

The latter relation is no longer understood in the sense that b is a function of a since both a and b are constants. In this case we should say " b is equal to the

value of the function f for the value of the independent variable equal to a "; for brevity it is usually said " b is equal to f of a ", but one should bear in mind the precise meaning of the statement indicated here. The notation $y|_{x=a}=b$ is also frequently used in such cases.

The point M of the graph of a function $y=f(x)$ corresponding to the abscissa $x=a$ has the ordinate $f(a)$, which is written as $M[a, f(a)]$ (Fig. 2).

10. Basic Elementary Functions. Composite Function.

Definition. The class of *basic elementary functions* includes the following functions:

- (1) the *power function* $y=x^a$ where the exponent a is real;
- (2) the *exponential function* $y=a^x$ where a is a positive number different from 1;
- (3) the *logarithmic function* $y=\log_a x$ to the base a which is a positive number different from 1;
- (4) the *trigonometric functions* $y=\sin x$, $y=\cos x$, $y=\tan x$ and also the less frequently used functions $y=\sec x$, $y=\csc x$ and $y=\cot x$;
- (5) the *inverse trigonometric functions* $y=\arcsin x$, $y=\arccos x$, $y=\arctan x$ and also $y=\operatorname{arcsec} x$, $y=\operatorname{arccsc} x$ and $y=\operatorname{arccot} x$.

These functions are well known from the elementary course of mathematics studied at schools, and we shall briefly discuss them in § 3.

The basic elementary functions can be used to construct more complicated functions by means of arithmetical operations (ad-

dition, subtraction, multiplication and division) and a new operation of forming "a function of a function" to which we proceed now.

Let a function $u = \varphi(x)$ of the argument x be defined on a set D , G being the range of that function u . Furthermore, let another function $y = f(u)$ be defined on the set G . Then, to each value of x belonging to the set D there corresponds a definite value of u belonging to the set G , and to u , in its turn, there corresponds a definite value of y . Therefore, ultimately, to each value of x from the set D there corresponds a definite value of y , and hence y is a function of x . Denoting this new function by $y = F(x)$ we can write down the expression of the function $F(x)$ in terms of the functions f and φ :

$$y = F(x) = f[\varphi(x)]$$

It is said that $F(x)$ as function of x is a *composite function*¹ ("a function of a function") formed of the functions f and φ . The function $u = \varphi(x)$ entering into the expression $f[\varphi(x)]$ is referred to as the *intermediate variable*.

Thus, the argument of a function can be an independent variable or an intermediate variable dependent on another independent variable. It is the latter case when a composite function appears. For instance, the function $y = \sin^2 x$ can be regarded as a composite function of x since y can be interpreted as a quadratic function of an intermediate variable which, in its turn, is equal to the sine of the independent variable x : if the intermediate variable is denoted by u we can write $y = u^2$ where $u = \sin x$.

A composite function can be obtained by applying the operation of forming a function of a function to more than two constituent functions. A given function can be represented as a composite one in many ways, and an appropriate choice of such a representation often facilitates the application of the techniques of mathematical analysis.

11. Elementary Functions.

I. The functions constructed of *basic* elementary functions and constants by means of a finite number of arithmetical operations and operations of forming a function of a function are called *elementary functions*.

For example, all the functions written below are elementary:

$$y = \frac{a^2}{a + 2b \tan^2 x} + 3c \sin^3 x, \quad y = \frac{2 + \sqrt[3]{x}}{3 - \sqrt{x}}$$

$$y = \log(x + \sqrt{1 + x^2}), \quad y = \frac{10^x - 1}{x^2}, \quad y = \arctan \sqrt{\frac{1 + \sin x}{1 - \sin x}}$$

¹ The term "composite function" characterizes only the way in which the analytical expression of the function is obtained and does not mean that this function necessarily has a complex structure in the ordinary sense.

The important role of the class of *elementary functions* lies in the fact that applications of mathematical analysis to problems of physics and engineering most often involve elementary functions.

A formula expressing an elementary function makes it possible to determine the domain of definition of the function, which is important for mathematical analysis. Let us consider some examples.

(1) The domain of definition of the function $y = x^2 + 2x - 5$ obviously coincides with the entire number scale.

(2) The domain of definition of the function $y = \frac{1}{x}$ is the whole number scale with the point $x = 0$ deleted for which this formula loses sense since the division by zero is impossible. Thus, the domain of definition consists of two infinite open intervals:

$$-\infty < x < 0 \quad \text{and} \quad 0 < x < +\infty$$

(3) The domain of definition of the function $y = \sqrt{1-x^2}$ is the closed interval $[-1, +1]$ since for every point of this interval, that is for every x satisfying the inequalities $-1 \leq x \leq 1$, the expression $\sqrt{1-x^2}$ takes on a real value while for any point lying outside this interval it does not assume a real value.

(4) The domain of definition of the function $y = \sqrt{x^2-1}$ is the union of the two infinite intervals $-\infty < x \leq -1$ and $+1 \leq x < +\infty$ (which can also be expressed by the inequality $|x| \geq 1$), that is the whole number scale with the interval $(-1, +1)$ deleted.

(5) The domain of definition of the function $y = \log x$ is the open infinite interval $0 < x < +\infty$.

(6) The domain of definition of the function $y = \frac{2x^2 - \log(x+5)}{\sqrt{8-x^2}}$ is the open interval $(-5, 2)$ since the expression $\log(x+5)$ is defined for $x > -5$ while the expression $\sqrt{8-x^2}$ for $x \leq 2$, and the point $x = 2$ should be excluded because the denominator of the fraction vanishes at that point.

In concrete problems we may encounter functions whose domains of definition are wider than the natural range of their independent variable coherent with the conditions of the problems. For example, the area S of a circle is expressed in terms of its radius r by the formula

$$S = \pi r^2$$

Although this formula is applicable to all real values of r , it is senseless to take a negative or zero value of the magnitude r since it is a radius of a circle which must be positive. The domain of definition of this function coherent with the geometrical conditions is restricted to the interval $0 < r < +\infty$.

Another example of this type is the functional relationship describing the basic law of the theory of elasticity known as Hooke's¹ law. It states that the elongation of a bar is directly proportional to the magnitude of the tensile force, that is $y = cx$ where y is the elongation, x is the magnitude of the force and c is a proportionality factor which is constant for the given bar. If this function is considered irrespective of the physical meaning of the magnitudes involved its domain of definition consists of all the real values of x . But, as is well known, Hooke's law only holds for small elongations, and this imposes restriction of the domain within which the formula remains applicable.

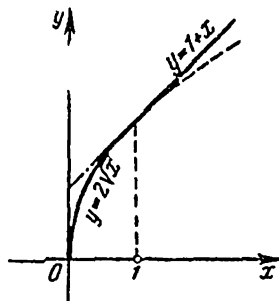


Fig. 3

We often deal with functions which have *different* analytical expressions on *different* parts of its domain of definition.

For instance, let us define a function, say $y = f(x)$, by means of the formulas

$$y = f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 2\sqrt{x} & \text{for } 0 \leq x \leq 1 \\ 1 + x & \text{for } x > 1 \end{cases}$$

This function is defined throughout the x -axis; its graph is depicted in Fig. 3.

Examples of *nonelementary* functions will be considered later on.

II. Algebraic and transcendental functions. Using the power function $y = x^n$ with a positive integral exponent n we form a *polynomial*

$$y = a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n.$$

The coefficients of the polynomial are arbitrary real numbers. If $a_0 \neq 0$ the function y is said to be a *polynomial of the n th degree*. A polynomial of *zeroth degree* is a constant. It is clear that every polynomial is a function over the whole number scale. A polynomial is also referred to as an *entire rational function*.

A *rational function* is a quotient of two polynomials:

$$y = \frac{a_0 x^n + a_1 x^{n-1} + \dots + a_n}{b_0 x^m + b_1 x^{m-1} + \dots + b_m}$$

The domain of definition of a rational function is the whole axis of abscissas with the exception of the points (the roots of the denominator) at which the denominator of the fraction vanishes.

¹ R. Hooke (1635-1703), an English physicist.

Here are examples of such functions:

$$y = x^3 - 3x + 5, \quad y = \frac{1}{x^2 + 1}, \quad y = \frac{\sqrt[3]{3x^2 - 1}}{x^3 + \pi}$$

When, for a given value of the independent variable x , the corresponding value of a rational function is computed, only arithmetical operations on the variable x and on the constant coefficients are performed, that is, addition, subtraction, multiplication (including raising to a positive power) and division. It should be noted that the coefficients of the polynomials in the numerator and in the denominator can be irrational numbers (for instance, this is the case in the last example).

A wider class of functions is obtained if, besides the arithmetical operations, the extraction of roots is allowed. For instance, we then can form the functions

$$y = \sqrt{x+1} + x^2, \quad y = \sqrt[3]{\sqrt{x+1}}, \quad y = \frac{\sqrt[3]{2x+1} + x}{x - \sqrt{x} + \pi}$$

The total number of operations must of course be finite. The functions obtained in this way enter into the class of *algebraic functions*.

The class of *algebraic functions* whose general definition will be given in Sec. 12 is not exhausted by the above functions.

Every function not belonging to the class of algebraic functions is called a *transcendental function*.

12. Implicit Functions. Multiple-valued Functions. Up to now we confined ourselves to those functions specified analytically for which the left member of the equality defining the function is the variable y *alone* while the right member is an expression only involving x . We shall call such functions *explicit*. Such are, for instance, all the functions discussed in Sec. 11.

But a more general equation connecting two variables and *not resolved* in one of them can also specify one variable as a function of the other.

For example, in the equation $2x + 3y = 6$ of a straight line the ordinate y can be regarded as a function of the abscissa x defined throughout Ox . Indeed, to every value of x there corresponds a definite value of y determined by this equation of the straight line. In such cases we say that the function is specified *implicitly*. To obtain its explicit expression we should *resolve* the equation in the dependent variable. In the above example this leads to the explicit function $y = 2 - \frac{2}{3}x$.

Now take the equation of the circle

$$x^2 + y^2 = 1$$

Here we see that to each value of x lying between -1 and $+1$ there correspond, instead of one, *two* values of y (see Fig. 4). Accordingly, equation $x^2 + y^2 = 1$ represents not one but *two* functions $y = +\sqrt{1-x^2}$ and $y = -\sqrt{1-x^2}$ each of which is defined on the interval $[-1, 1]$. If x assumes a value not belonging to this interval the equation from which y is found has no real solutions.

We can also say that the equation $x^2 + y^2 = 1$ defines y as a *multiple-valued* (in this particular case two-valued) function of x . The functions $y = +\sqrt{1-x^2}$ and $y = -\sqrt{1-x^2}$ are then called the *single-valued branches of the multiple-valued function*¹. The general interpretation of a functional relationship including multiple-valued functions considered as aggregates of single-valued branches will be often used in what follows. In this approach we can say that an *implicit function* y (which can be multiple-valued) of an *independent variable* x is a function whose values are found from an equation connecting x and y and not solved in y .

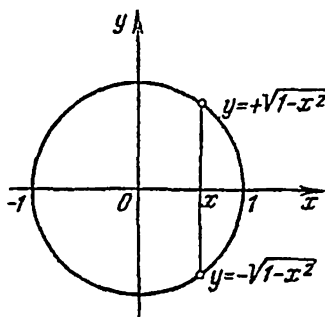


Fig. 4

To obtain the explicit representation of an implicit function the given equation must be resolved with respect to y . This is by far not always readily achieved and sometimes is even impossible. For instance, the equation

$$y^3 - 3xy + x^3 = 0$$

specifies y as a function of x , but it is more difficult than in the above examples to obtain its explicit expression since this is connected with the solution of the algebraic equation of the third degree. If we take the equation

$$y + x2^y = 1$$

which also determines y as a function of x we see that it cannot be solved algebraically, and y cannot be expressed explicitly in terms of x . In this case it is rather difficult to analyse the domain of definition of the function y and separate its single-valued branches.

¹ We do not consider functions which can be obtained from the same equation such that for some x 's their values are given by the formula $y = +\sqrt{1-x^2}$ and for the other x 's by the formula $y = -\sqrt{1-x^2}$. The graph of such a function is not a continuous curve; for instance, let the reader think about the graph of the function specified by the conditions $y = +\sqrt{1-x^2}$ for $x \in [-1, 0]$ and $y = -\sqrt{1-x^2}$ for $x \in [0, 1]$.

Equations of this type can be solved *numerically*. Numerical methods which make it possible to compute an approximation to the value of y corresponding to a given particular value of x will be discussed later on (see Sec. 77).

One must not think that any equation connecting x and y necessarily defines a function of one variable. For example, there are no *real* values of x and y satisfying the equation $x^2 + y^2 + 1 = 0$ since its left member cannot be equal to zero. A more detailed analysis of implicit functions will be given in Sec. 117.

The notion of an implicit function makes it possible to state the general definition of an algebraic function.

Definition. An *algebraic function* is a function $y = y(x)$ specified by an algebraic equation

$$P_0(x)y^n + P_1(x)y^{n-1} + \dots + P_n(x) = 0$$

where $P_j(x)$ ($j=0, 1, \dots, n$) are polynomials.

In the general case an algebraic function is multiple-valued, and it is more natural to consider it as being defined in a domain of the complex plane.

§ 3. Characteristics of Behaviour of Functions. Some Important Examples

13. Simplest Characteristics of Functions. To investigate a function means to characterize its *variation* (or, as we say, its *behaviour*) as the independent variable changes. In what follows we shall always suppose (unless the contrary is stated) that *when an independent variable changes it increases* and that *when it passes from smaller values to greater ones it assumes all the intermediate values*.

As the reader proceeds with the study of this course, and more sophisticated methods become available we shall pass to a more complete description of functions. But now, since we can only use the techniques of elementary mathematics and fundamentals of analytic geometry the behaviour of a function $y = f(x)$ is characterized by the following simplest features:

I. *Zeros of a function and its sign in the given interval.* II. *The property of being even or odd.* III. *Periodicity.* IV. *Increase or decrease in the given interval.*

I. **Definition.** A value of x for which a function vanishes, that is $f(x) = 0$, is called a *zero* (or a *root*) of the function.

On an interval where a function is positive its graph lies above the x -axis, and on an interval of negative sign under the x -axis. At a zero the graph of the function has a common point with Ox (Fig. 5).

II. **Definition.** A function $y = f(x)$ is said to be *even* if the change of the sign of any value of the independent variable does

not affect the value of the function:

$$f(-x) = f(x)$$

A function $y = f(x)$ is called *odd* if the change of the sign of any value of the argument results in the change of the sign of the function while its absolute value remains invariable:

$$f(-x) = -f(x)$$

In this definition it is assumed that the function is defined in a domain symmetric about the origin of coordinates.

Examples of even functions are $y = x^2$, $y = \cos x$ and examples of odd ones are $y = x^3$, $y = \sin x$, $y = \tan x$.

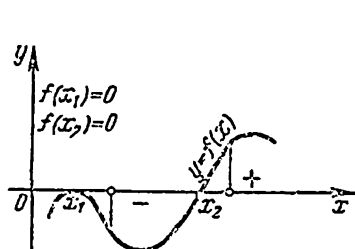


Fig. 5

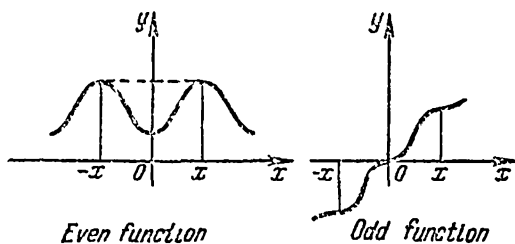


Fig. 6

The graph of an even function is symmetric with respect to Oy . The graph of an odd function is symmetric about the origin. Let the reader prove these properties (see Fig. 6).

An arbitrary function can of course be neither even nor odd; for example, such are the functions $y = x + 1$, $y = 2^x$ and $y = 2 \sin x + 3 \cos x$.

III. Definition. A function $y = f(x)$ is said to be *periodic* if there exists a number $a \neq 0$ such that for any x belonging to the domain of definition of the function the values $x + a$ and $x - a$ of the independent variable also belong to the domain of definition, and the identity

$$f(x + a) = f(x)$$

holds. Every such number a is called a *period* of the function.

If a function $f(x)$ is periodic with period a the equalities

$$\begin{aligned} f(x + 2a) &= f(x), & f(x + 3a) &= f(x), \\ f(x - a) &= f(x), & f(x - 2a) &= f(x) \end{aligned}$$

take place, and, generally

$$f(x + ka) = f(x)$$

for any x and any (positive or negative) integer k . Thus, if a is a period the number ka is also a period.

In what follows, when speaking about a period of a function we shall mean, as a rule, its *least positive period* (referred to as the *primitive period*) provided that it exists.

The function $y = \sin x$ is an example of a periodic function; its period is equal to 2π . If a periodic function is not defined at a point x_0 the same applies to all the points $x_0 + ka$ where a is the period and k any integer. For instance, the function $y = \tan x$ whose period is equal to π is not defined at the points $\frac{\pi}{2} + k\pi$.

The function $y = \sqrt{\sin x}$ is periodic with period 2π . It is defined on the intervals where $\sin x \geq 0$, that is for $2k\pi \leq x \leq (2k+1)\pi$ where k is an integer (which may be positive, negative or zero).

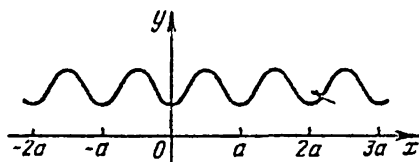


Fig. 7

If a periodic function is defined throughout the number scale it is sufficient to investigate it in any interval whose length is equal to the period of the function, for instance, in the

interval $0 \leq x < a$ (a is the period), since the values of the function at all the other points of Ox are found by means of simple shifts, along the axis of abscissas, of the points of the graph of the function on the interval $0 \leq x < a$, by ka where $k = 0, \pm 1, \pm 2, \dots$. Thus, the whole graph of a periodic function with period a is obtained from its part on an interval whose length is equal to the period of the function by an infinite number of shifts by a along Ox in the positive and negative directions (see Fig. 7).

IV. Important characteristics of the behaviour of a function are its *increase* and *decrease*.

Definition. A function is said to be *increasing* on an interval if to greater values of the argument belonging to that interval there correspond greater values of the function; similarly, it is called *decreasing* if to greater values of the argument there correspond smaller values of the function.

Thus, $f(x)$ *increases on the interval* $[a, b]$ if for any values x_1 and x_2 of the argument satisfying the condition $a \leq x_1 < x_2 \leq b$ the inequality $f(x_1) < f(x_2)$ holds; similarly, it *decreases on* $[a, b]$ if for any x_1 and x_2 satisfying the condition $a \leq x_1 < x_2 \leq b$ the inequality $f(x_1) > f(x_2)$ takes place.

For example, the function $y = 2^x$ increases on the entire x -axis while the function $y = 2^{-x}$ decreases on it. The function $y = x^2$ decreases on the interval $(-\infty, 0]$ and increases on the interval $[0, \infty)$.

If the graph of a function is traced from left to right (this corresponds to the increase of the argument x), then for an increasing function the moving point of its graph goes upward

(relative to the positive direction of Oy), and for a decreasing function it moves downward (see, respectively, the lines AB and CD in Fig. 8).

The increase and the decrease of a function can be interpreted in a broader sense. A function $f(x)$ is called *nondecreasing* on an interval $[a, b]$ (or *increasing in the broad sense*) if the condition $a \leq x_1 < x_2 \leq b$ implies the slack (nonstrict) inequality $f(x_1) \leq f(x_2)$. Accordingly, if this implies $f(x_1) \geq f(x_2)$ the function is said to be *nonincreasing* (or *decreasing in the broad sense*).

This type of increase or decrease in the broad sense is most often characteristic of functions having different analytical expressions on different intervals; for instance, the function whose graph is shown in Fig. 3 is nondecreasing throughout Ox .

An interval of variation of the independent variable on which the function increases is called its *interval of increase*, and an interval on which the function decreases is called its *interval of decrease*. Both types of intervals are referred to as *intervals of monotonicity* of the function, and the function itself is said to be *monotone* in such an interval.

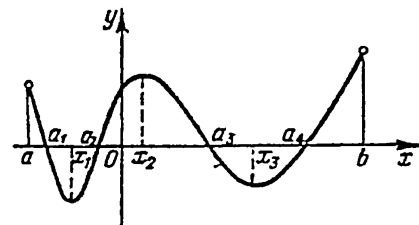


Fig. 9

It is usually possible to break up the interval on which a function is considered into a number of subintervals on each of which the function is monotone. For instance, see Fig. 9 where the graph of a function defined on the interval $[a, b]$ is shown; this interval is split into the intervals $[a, x_1]$, $[x_1, x_2]$, $[x_2, x_3]$ and $[x_3, b]$, on which, respectively, the function decreases, increases, decreases and increases.

For simple functions whose graphs are known the intervals of monotonicity are easily determined; later on (see Sec. 59) some techniques will be developed which make it possible to find the intervals of monotonicity of a function without constructing its graph.

Another important characteristic of a function is given by the following

Definition. If a function assumes, on a given interval, a value which is greater (less) than all the other values, this value is called the *greatest* (the *least*) value of the function on that interval.

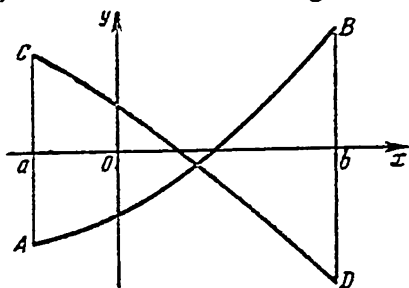


Fig. 8

greater the number of points M constructed in the coordinate plane, the more accurate the representation of the graph.

Knowing the graph of a function $y=f(x)$ we can readily construct the graphs of the functions $y=f(x)+b$, $y=f(x+a)$, $y=Af(x)$ and $y=f(kx)$ where b , a , A and k are given numbers. The points of the graph of the function $y=f(x)+b$ are obtained by the translation of the points of the graph of $y=f(x)$ along the y -axis: the points are moved by a distance of $|b|$ upward if $b>0$ and downward if $b<0$. The whole graph thus undergoes a parallel displacement upward or downward by b units of length.

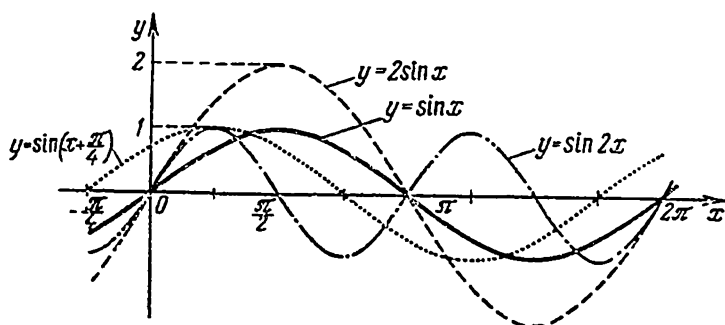


Fig. 10

For constructing the graph of $y=f(x+a)$ we perform a parallel displacement of the graph of the original function $f(x)$ along Ox shifting it by a distance of a to the left if $a>0$ and by $|a|$ to the right if $a<0$. This follows from the fact that the values of the functions $y=f(x)$ and $y=f(x+a)$ are equal when the value of x is equal to x_0 for the former function and to x_0-a for the latter (x_0 is arbitrary).

To obtain the graph of the function $y=Af(x)$ we multiply all the ordinates of the original graph of A . For a positive A this means that the ordinates are increased A times. If $A<0$ we construct the graph of the function $y=|A|f(x)$ and then take its reflection through Ox . If $k>0$ the graph of the function $y=f(kx)$ is obtained by decreasing k times the abscissas of all the points of the graph of $y=f(x)$. If $k<0$ we construct the graph of $y=f(|k|x)$ and take its reflection in the y -axis.

Finally, to obtain the graph of the function $y=Af[k(x+a)]+b$ we construct, in succession, the graphs of the functions $f_1(x)=f(x+a)$ and $f_2(x)=f[k(x+a)]$ (resulting, respectively, from a parallel displacement along Ox and uniform $|k|$ -fold extension (contraction) along Ox with centre at the point $x=-a$), and then the graphs of $f_3(x)=Af_2(x)$ and $y=f_3(x)+b$ (obtained by an extension and a translation along Oy). See Fig. 10 which

demonstrates the construction of the graphs of the functions $y = 2 \sin x$, $y = \sin 2x$ and $y = \sin\left(x + \frac{\pi}{4}\right)$ from the given graph of the function $y = \sin x$.

Knowing the graphs of two functions $y = f_1(x)$ and $y = f_2(x)$ we can readily construct, by purely geometrical means, the graph of the function $y = a_1 f_1(x) + a_2 f_2(x)$ where a_1 and a_2 are constants. To this end, we multiply the ordinates of the functions $f_1(x)$ and $f_2(x)$ by, respectively, a_1 and a_2 , and then add together the ordinates of the curves thus obtained. An expression of the form $a_1 f_1(x) + a_2 f_2(x)$ is called a *linear combination* of the given functions $f_1(x)$ and $f_2(x)$. A linear combination of an arbitrary number of functions is formed in the same manner.

15. Direct Variation and Linear Function. Increment. When two variables are related so that their ratio remains constant *one of them is said to vary directly as the other*; or such variables are said to be (directly) *proportional*. The fact that y is proportional to x is expressed by the formula

$$y = ax$$

where $a \neq 0$ is a constant referred to as the *constant of proportionality* or *factor of proportionality* or *constant of variation*.

A feature of this functional relationship is that if x_1, y_1 and x_2, y_2 are two pairs of mutually corresponding values of x and y then $y_1 : y_2 = x_1 : x_2$.

If $y = ax$ then $x = \frac{1}{a}y$, and hence if y is proportional to x then x also varies directly as y , but the new proportionality factor is the *reciprocal* of the former.

Direct variation is a particular case of a more general functional relationship described by a *linear function* of the form

$$y = ax + b$$

where a and b are constants.

If $b = 0$ and $a \neq 0$ the linear functional relationship reduces to a direct variation; if $b \neq 0$ the variable y is no longer proportional to x .

Every linear function is defined throughout the number scale. As is known from analytic geometry, the graph of a linear function is a *straight line*. The constant coefficients a and b are, respectively, equal to the slope of the straight line (that is to the tangent of the angle of inclination of the line to Ox) and to the ordinate of the point of the line with abscissa $x = 0$. For $b = 0$ the graph of the linear function (expressing a direct variation) is a straight line passing through the origin. It is clear that for a positive a the linear function is *increasing* on the entire x -axis and for $a < 0$ it is *decreasing* throughout Ox .

Now we introduce an important definition which, together with the corresponding notation, will be permanently used.

Definition. Let a variable u pass from its arbitrary (initial) value u_1 to another (terminal) value u_2 . The difference between the terminal and the initial values is called the *increment* of the variable u and is denoted by Δu :

$$\Delta u = u_2 - u_1$$

The increment Δu can be positive, negative or zero. Accordingly, in the first case the variable u increases when it passes from u_1 to $u_2 = u_1 + \Delta u$, in the second case it decreases and in the third case, when $u_1 = u_2$, it remains invariable: $\Delta u = 0$.

Note that the expression Δu should not be thought of as the product of Δ by u : this is a unified symbol designating the difference between the terminal and the initial values of the variable u .

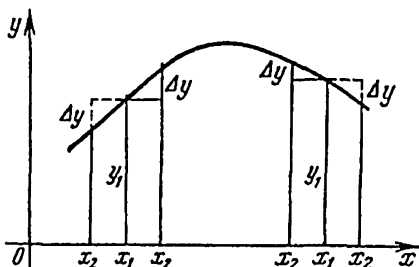


Fig. 11

If the independent variable x receives an increment Δx the function $y = f(x)$ also gains an increment: $\Delta y = f(x + \Delta x) - f(x)$. On the graph the increment of the function is represented by a line segment parallel to Oy taken with the sign $+$ or $-$ depending on whether the ordinate increases or decreases as the abscissa passes from the initial value to the terminal value (see Fig. 11 where the initial values of the abscissa are denoted as x_1 and the terminal ones as x_2). For an arbitrary function its increment depends not only on the increment Δx of the argument but also on the chosen initial value of x ; this is clearly seen in Fig. 11.

Now we proceed to prove the following important property of the linear function.

Direct Theorem. The increment of a linear function is directly proportional to the increment of the argument and does not depend on the initial value of the argument.

Proof. Let the argument receive an increment Δx and pass from a value x of the abscissa to the value $x + \Delta x$; then the function gains an increment Δy and passes from its value y to the new value $y + \Delta y$. Hence, we have

$$y = ax + b$$

and

$$y + \Delta y = a(x + \Delta x) + b$$

On subtracting the former equality from the latter we obtain

$$\Delta y = a\Delta x$$

which is what we set out to prove.

This relation can also be derived geometrically. Indeed, it is seen from Fig. 12 that Δy and Δx are the lengths of the legs of the right triangle ABC , and $\tan \alpha = a$. Therefore

$$\Delta y = \tan \alpha \cdot \Delta x = a \Delta x$$

This property completely characterizes the linear function in the sense that there are no other functions distinct from linear ones possessing the same property. To show this we shall prove the converse of the above theorem.

Reciprocal Theorem. If the increment of a function is directly proportional to that of the argument and is independent of its initial value the function is linear.

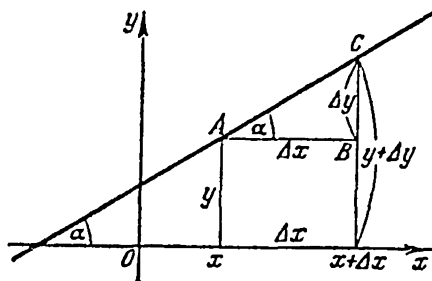


Fig. 12

Proof. Suppose that we are given a function $y = f(x)$ for which the increment Δy taken for any x is proportional to the increment of the independent variable Δx :

$$\Delta y = a \Delta x$$

where a is a constant. Let the function $f(x)$ assume a value b for $x=0$, that is $f(0)=b$. If the argument passes to an arbitrary new value x the function takes on a new value $y=f(x)$. In this case the increment of the argument is equal to $\Delta x = x - 0 = x$ while the increment of the function is $\Delta y = y - b$. Consequently, by the hypothesis, we must have $y - b = ax$ whence

$$y = ax + b$$

for any x , which is what we wanted to prove.

Every *uniform process* in which to any two equal increments of one variable there correspond equal increments of another variable is described by a linear function.

For instance, a motion is uniform if during any equal intervals of time (t) equal path lengths (s) are travelled; then $\Delta s = v \Delta t$ where v is the constant velocity of motion. The basic property of the linear function then implies that $s = vt + s_0$ where s_0 is the path corresponding to the time moment $t=0$. This is the *equation of uniform motion*.

Uniform processes are the simplest but they play a very important role.

16. Quadratic Function. A *quadratic function* is specified by the equality

$$y = ax^2 + bx + c$$

where a , b and c are constants ($a \neq 0$). Every quadratic function is defined throughout the x -axis.

In the simplest case when $b=c=0$ the quadratic function has the form $y=ax^2$, its graph being a parabola with vertex at the origin and axis coinciding with the positive y -axis for $a > 0$ and with the negative y -axis for $a < 0$. As is known from analytic geometry, in the general case the graph of the function $y=ax^2+bx+c$ is a parabola with axis parallel to Oy and vertex at the point M_0 with abscissa $x=-\frac{b}{2a}$ and ordinate $y=-\frac{b^2-4ac}{4a}$. The parabola opens upwards or downwards depending on whether $a > 0$ or $a < 0$ (Fig. 13).

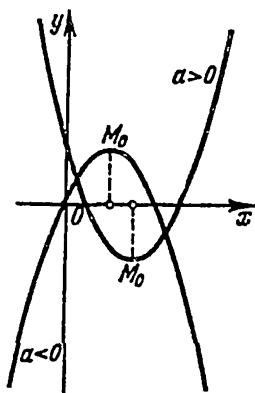


Fig. 13

The quadratic function has only one change in the character of its variation: it either first increases and then decreases or vice versa. Namely, if $a > 0$ it decreases on the interval $(-\infty, -\frac{b}{2a})$, attains the least value $y = -\frac{b^2-4ac}{4a}$ at the point $x = -\frac{b}{2a}$

and then increases on the interval $(-\frac{b}{2a}, +\infty)$. In this case the greatest value does not exist. Conversely, for $a < 0$ the function has no least value; it first increases and attains its greatest value at the point $x = -\frac{b}{2a}$ and then decreases.

For example, the quadratic function $y=x^2-5x+4$ has $a=1 > 0$ and therefore possesses the least value. There is no need in constructing its graph to determine its least value since the above formulas indicate that it is assumed at the point $x = -\frac{b}{2a} = 2.5$ and is equal to $y = -\frac{b^2-4ac}{4a} = -2.25$. (The least value $y = -2.25$ can also be obtained by substituting $x=2.5$ into the expression of the function.)

Problems of determining the greatest and the least values of functions often arise in various applications. Let us consider two examples.

(1) It is required to find, among all the rectangles with a given perimeter P , the one of the greatest area.

Denote one of the sides of the rectangle by x ; then the other side is equal to $\frac{P}{2} - x$. Consequently, the area S of the rectangle

is determined by the formula

$$S = x \left(\frac{P}{2} - x \right) = \frac{P}{2}x - x^2$$

and thus is a quadratic function of x . Since $a = -1 < 0$ the function possesses the greatest value; it is attained for $x = -\frac{b}{2a} = \frac{P}{4}$ and is equal to

$$\frac{P}{2} \cdot \frac{P}{4} - \left(\frac{P}{4} \right)^2 = \frac{P^2}{16}$$

But a rectangle with perimeter P whose one side has the length $\frac{P}{4}$ is a square. Hence, from all the rectangles with given perimeter the square possesses the greatest area.

(2) A shell is shot vertically upward from an anti-aircraft gun placed at the height $h_0 = 200$ m above sea-level, the initial velocity of the shell being $v_0 = 500$ m/s. How long does it take the shell to reach the maximum height above sea-level and what is this maximum height (air resistance is neglected)?

Air resistance being neglected, the motion of the shell can be considered as uniformly decelerated with negative acceleration $-g$. Then, as is known, the height h of the shell above sea-level at time t is expressed by the formula

$$h = h_0 + v_0 t - \frac{g}{2} t^2$$

Hence, h is a quadratic function of t , and $a = -\frac{g}{2} < 0$. The greatest value of the function h (that is the maximum height of the shell above sea-level) is attained for

$$t = -\frac{b}{2a} = \frac{v_0}{g} = \frac{500}{9.8} \approx 51 \text{ s}$$

Consequently, the shell reaches its maximum height 51 s after the shot, and the greatest value of h is

$$H = 200 + 500 \cdot 51 - \frac{9.8}{2} \cdot 51^2 = 12,955 \text{ m} \approx 13 \text{ km}$$

In reality, due to air resistance, H is considerably smaller.

17. Inverse Variation. Linear-fractional Function. When the ratio of one variable to the reciprocal of another variable is constant *one of them is said to vary inversely as the other*. The inverse variation of y as function of x is described by the formula

$$y = \frac{a}{x}$$

where a is a constant ($a \neq 0$). We also say that y is *inversely proportional to x* .

A feature of this functional relationship is that if x_1, y_1 and x_2, y_2 are two pairs of mutually corresponding values of x and y then $y_2:y_1 = x_2:x_1$.

If $y = \frac{a}{x}$ then $x = \frac{a}{y}$; hence if y is inversely proportional to x then x , in its turn, is inversely proportional to y with the same coefficient a .

Inverse variation is a particular case of a more general functional relationship described by a *linear-fractional* function of the form

$$y = \frac{ax+b}{cx+d} \quad (*)$$

where a, b, c and d are constants ($c \neq 0$). For $a=0$ and $d=0$ it reduces to inverse proportionality.

Linear-fractional function (*) is defined over the whole x -axis except for the point $x = -\frac{d}{c}$ (which coincides with the point $x=0$ in the case of the function $y = \frac{a}{x}$).

We shall suppose that $bc - ad \neq 0$ since, if otherwise, the linear-fractional function degenerates into a constant. For if $bc = ad$, we put $a = cm$ and $b = \frac{ad}{c} = dm$ where $\frac{a}{c} = m$, and obtain

$$y = \frac{ax+b}{cx+d} = \frac{cmx+dm}{cx+d} = m \frac{cx+d}{cx+d} = m$$

Performing a translation of the coordinate axes we readily prove that the graph of a linear-fractional function is an equilateral hyperbola with asymptotes parallel to the coordinate axes. Indeed, if the origin of the translated system is placed at the point $O'(-\frac{d}{c}, \frac{a}{c})$ the equation of the graph of linear fractional function (*) in the coordinates $O'x'y'$ takes the form

$$y' = \frac{a'}{x'} \quad \text{where} \quad a' = \frac{bc-ad}{c^2}$$

(by the hypothesis, $bc - ad \neq 0$, and therefore $a' \neq 0$). The latter equation describes an equilateral hyperbola whose asymptotes are the new axes $O'x'$ and $O'y'$, that is the straight lines

$$x = -\frac{d}{c} \quad \text{and} \quad y = \frac{a}{c}$$

Thus, to construct the graph of a linear-fractional function we draw through the point $(-\frac{d}{c}, \frac{a}{c})$ the straight lines parallel to the coordinate axes, take them as the new coordinate axes and trace the equilateral hyperbola with semi-axes equal to $\sqrt{2|a'|}$.

The sign of the expression $a' = \frac{bc-ad}{c^2}$, i.e. the sign of $bc-ad$, determines the quadrants in which the hyperbola lies: if $bc-ad > 0$ the hyperbola is in the first and third quadrants and if $bc-ad < 0$ it is in the second and fourth quadrants.

As an example, consider the function

$$y = \frac{x+1}{x-1}$$

Here we have $a=1$, $b=1$, $c=1$ and $d=-1$, and, consequently, the asymptotes of the hyperbola are the lines $x = -\frac{d}{c} = 1$ and $y = \frac{a}{c} = 1$. Furthermore, $a' = \frac{bc-ad}{c^2} = 2$, and hence the semi-axes of the hyperbola are equal to $\sqrt{2|a'|} = 2$. Fig. 14 demonstrates

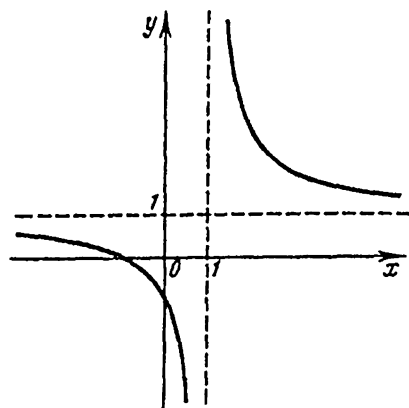


Fig. 14

the graph of this function; it lies in the first and third quadrants formed by the asymptotes since $bc-ad = 2 > 0$.

In the particular case of inverse proportionality, i.e. for $y = \frac{a}{x}$, the asymptotes coincide with the coordinate axes.

The behaviour of the linear-fractional function is clearly seen from its graph:

In every interval of the x-axis not containing the point $x = -\frac{d}{c}$ linear-fractional function () either only increases or only decreases.*

The character of variation (that is increase or decrease) depends on the sign of the expression $bc-ad$: if $bc-ad > 0$ the function is decreasing and if $bc-ad < 0$ it is increasing.

Let us again consider the function $y = \frac{x+1}{x-1}$ whose graph is depicted in Fig. 14. For the values of x less than 1 the function permanently remains less than unity and decreases as x approaches unity. When the argument x passes from the values of x smaller than 1 to those greater than 1 the function "makes an infinite jump" and passes from the values which are negative and arbitrarily large in their absolute values to arbitrarily large positive values. After this jump, as x increases, the function decreases remaining greater than unity and approaching unity when x grows indefinitely.

§ 4. Inverse Function. Power, Exponential and Logarithmic Functions

18. Inverse Function. Consider a function $y=f(x)$ with a domain of definition (that is the admissible set of the values of x) D and a range (the set of the corresponding values of y) G . We say that the function $y=f(x)$ specifies a *mapping*¹ of the set D onto the set G . The definition of the notion of a function implies that to each point of the set D there corresponds a uniquely determined point of the set G . If, in this mapping, each point of the set G corresponds to a single point of the set D the *mapping* (the *correspondence*) is said to be *one-to-one*. For a one-to-one mapping to different points of the set D there correspond *different* points of the set G .

For instance, the function $y=x^3$ specifies a one-to-one mapping of the whole axis Ox onto Oy since to every value of x there corresponds a single value of y , and, conversely, to each y there corresponds a single value of x equal to $\sqrt[3]{y}$.

For the function $y=x^2$ the mapping of Ox (which is its domain of definition) onto the interval $[0, \infty)$ of the y -axis (which is the range of the function) is not one-to-one because to every $y > 0$ there correspond two values of x : $x = +\sqrt{y}$ and $x = -\sqrt{y}$.

The function $y=\sin x$ maps the entire Ox axis onto the interval $[-1, 1]$ of the y -axis. It is apparent that this mapping is not one-to-one: to every value of y belonging to the interval $[-1, 1]$ there correspond an infinite number of values of x .

If a function $y=f(x)$ performs a one-to-one mapping of a set D onto a set G then, according to the above definition, this correspondence assigns to every value of y belonging to the set G a single value of x belonging to the set D . Therefore we can say that this defines a function $x=\varphi(y)$ whose domain of definition is G and range D . We see that in this construction the roles of the variables x and y are interchanged: y becomes an independent variable and x a function. The functions $y=f(x)$ and $x=\varphi(y)$ thus related are said to be *mutually inverse*.

If the mapping of the set D onto the set G is not one-to-one then, to some values of y (and possibly to all of them) there correspond several different values of x , and we thus encounter the situation which was described in Sec. 12 where implicit functions were considered. Namely, in this case the inverse function $x=\varphi(y)$ is *multiple-valued* (we also say that the *inverse mapping* specified by the function $y=f(x)$ is *many-to-one*). This question will be discussed in more detail later on, and now we proceed

¹ In modern mathematics the terms "function" and "mapping" and also "transformation" and "operator", are used synonymously—Tr.

to establish a geometrical relationship between the graphs of mutually inverse functions and to study techniques for constructing the graph of the inverse of a given function.

Let a variable y be represented as an implicit function of x :

$$y = f(x) \quad (*)$$

The same equation can be interpreted as defining x as an *implicit* function of y . To find the explicit expression of this function we should solve (provided this is possible) equation (*) with respect to x :

$$x = \varphi(y) \quad (**)$$

It is clear that if the correspondence between the points of the ranges of the functions x and y is one-to-one the functions $y = f(x)$

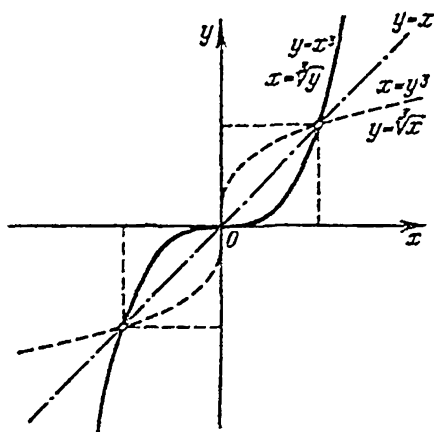


Fig. 15

and $x = \varphi(y)$ are single-valued and mutually inverse. For instance, if $y = x^3$ then $x = \sqrt[3]{y}$, and if $y = 10^x$ then $x = \log y$ (in both cases there holds the condition that the correspondence is one-to-one). It is easily verified that the inverse of a linear function is also a linear function and that the inverse of a linear-fractional function is also a linear-fractional function.

If a function is originally specified implicitly, i.e. by an equation connecting x and y , the resolution of the equation with

respect to y and with respect to x results in a pair of mutually inverse functions (of course, on condition that the correspondence between the points of the ranges of x and y is one-to-one). For example, the equation $xy + 3x - y - 1 = 0$ defines the mutually inverse functions $y = \frac{3x-1}{-x+1}$ and $x = \frac{y+1}{y+3}$ (both functions are linear-fractional).

Equations (*) and (**) describe one and the same curve in the xy -plane, but for the function f the axis of its argument is x while for the function φ this role is played by Oy .

For instance, if $y = x^3$ then $x = \sqrt[3]{y}$. The graph of these relationships is a *cubical parabola* (Fig. 15).

If we want to construct the graphs of two mutually inverse functions so that x is the axis of the argument for both of them we should denote the independent variable in formula (**) by x

and the function by y , that is, like in formula (*). Then we receive

$$y = \varphi(x)$$

In this notation the letter x designates the independent variable and the letter y the function for both mutually inverse functions $y = f(x)$ and $y = \varphi(x)$. For example, the functions $y = x^3$ and $y = \sqrt[3]{x}$ or $y = 10^x$ and $y = \log x$ are mutually inverse.

Note that if the operations indicated by two mutually inverse functions are performed, in succession, on an arbitrary value of x , this value remains invariable. For example, $\sqrt[3]{x^3} = x$, $(\sqrt[3]{x})^3 = x$ and $\log 10^x = x$, $10^{\log x} = x$.

In the general form this property can be written as

$$\varphi[f(x)] = x \text{ and } f[\varphi(x)] = x$$

that is, when applied in succession, the symbols of two mutually inverse functions "cancel out". In the first relation the value of x should belong to the domain of definition of the function $y = f(x)$ and in the second relation to the domain of definition of $y = \varphi(x)$. For instance, the identity $\log 10^x = x$ is valid for all x while the identity $10^{\log x} = x$ only holds for $x > 0$.

If the correspondence between the points of the ranges of x and y is not one-to-one these relations can be violated; for instance, this is the case for $\sqrt{x^2}$ which has the two values x and $-x$.

There is a simple relationship between the graphs of $y = f(x)$ and $y = \varphi(x)$:

The graphs of two mutually inverse functions $y = \varphi(x)$ and $y = f(x)$ are symmetric with respect to the bisector of the first and third quadrants.

Indeed, let

$$y = f(a) = b$$

for $x = a$. The point $M(a, b)$ belongs to the graph of the given function. Therefore the point M' with coordinates (b, a) belongs to the graph of the inverse function since $a = \varphi(b)$. We have (see Fig. 16) $\triangle ON'M' = \triangle ONM$, which implies $\angle N'OM' = \angle NOM$ and $OM' = OM$. It follows that, in the first place, the bisector OR of the first and third quadrants serves as the bisector of the angle MOM' as well, and, in the second place, the triangle $\triangle MOM'$ is isosceles. Therefore, the bisector OR is in fact a symmetry axis of the triangle MOM' . Thus, for every point of the graph of the given function there is a point on the graph of

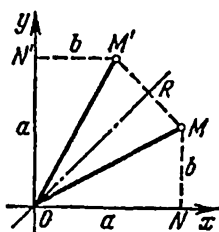


Fig. 16

the inverse function such that these points are symmetric about the bisector of the first and third quadrants. Since the converse is, apparently, also true the desired assertion has thus been proved.

Consequently, if the graph of a function is given the graph of its inverse can be obtained from the former by a mere reflection in the bisector of the first and third quadrants, that is, by the rotation about the bisector through 180° . Such a construction is illustrated in Fig. 15 where it is shown how the graph of the

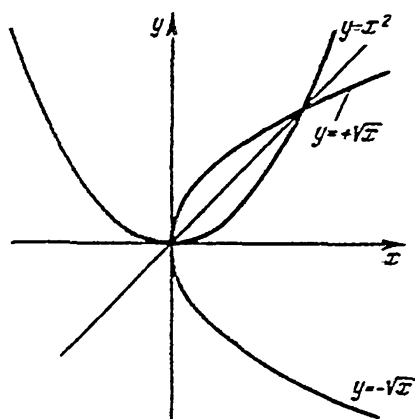


Fig. 17

function $y = \sqrt[3]{x}$ is obtained from the graph of its inverse $y = x^3$.

Now we proceed to establish conditions under which a mapping of a set of values of x onto a set of values of y is one-to-one; these conditions guarantee the existence of the single-valued inverse of the function $y = f(x)$. If to every value of x there corresponds a single value of y , every straight line parallel to the axis of ordinates intersects the graph of the function at no more than one point, and vice versa.

If the correspondence attributing certain values of y to the values of x possesses the same property, the above condition should also hold for the straight lines parallel to the axis of abscissas. It is clear, geometrically, that *these conditions hold if the given function is monotone; then the inverse function is also monotone.*

Here we shall not present a rigorous proof of this assertion limiting ourselves to its geometrical visuality implied by the relationship between the graphs of two mutually inverse functions established above.

If the given function is not monotone the situation becomes more complicated. For instance, take the function $y = x^2$ whose graph is a parabola (see Fig. 17). Let us split the domain of definition of the function (which is the whole number scale) into two intervals on which the function is monotone: $(-\infty, 0]$ and $[0, \infty)$. In each of these intervals the function $y = x^2$ possesses the inverse: for the interval $(-\infty, 0]$ the inverse function is $y = -\sqrt{x}$ and for $[0, \infty)$ it is $y = +\sqrt{x}$ (Fig. 17). The symbol $\sqrt{x}$ designates the arithmetic (i.e. positive) root.

As in Sec. 12, we can say that the inverse of the function $y = x^2$ is a *multiple-valued* (in this concrete case two-valued) function, the functions $y = +\sqrt{x}$ and $y = -\sqrt{x}$ being its *single-*

valued branches. (In future we shall write, for brevity, $y = \sqrt{x}$ instead of $y = +\sqrt{x}$.)

In what follows, when speaking of mutually inverse functions, we shall assume, without further stipulation, that they are both monotone in their domains of definition.

If a function in question is not monotone we break up its domain of definition, as in the above example, into intervals of monotonicity and consider the corresponding single-valued branch of the inverse function for every such interval. Further examples will be considered in Sec. 22 in connection with inverse trigonometric functions.

19. Power Function. The power function

$$y = x^\alpha$$

is defined throughout the x -axis for positive integral values of the exponent α . The shapes of the graphs of the functions $y = x^n$ for various positive integers n are demonstrated in Fig. 18. For even (odd) n these functions are even (odd), and therefore it is sufficient to investigate their general properties only for $x \geq 0$.

The greater n , the closer to the axis of abscissas lie the corresponding graphs within the interval $(0, 1)$ and the faster the moving points of the graphs go upward as x approaches unity.

Applying the results of Sec. 18 we can immediately construct

the graphs of the inverse functions $y = \sqrt[n]{x} = x^{\frac{1}{n}}$ of the functions $y = x^n$. In the construction it should be taken into account that

for an odd n both functions $y = \sqrt[n]{x} = x^{\frac{1}{n}}$ and $y = x^n$ are defined over the whole number scale, are monotone, and the shape of their graphs resembles Fig. 15 (which corresponds to the case $n=3$). If n is even the function $y = x^n$ is again defined throughout Ox while the inverse function splits into two branches $y = +\sqrt[n]{x}$ and $y = -\sqrt[n]{x}$ having the interval $[0, \infty)$ as their common domain of definition. The graphs of these functions are similar to those shown in Fig. 17 (corresponding to the case $n=2$).

Now let us demonstrate by examples the properties of the power function for nonintegral positive rational values of the exponent.

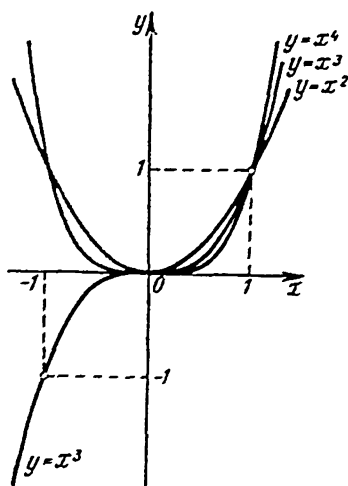


Fig. 18

The function $y = x^{\frac{2}{3}} = \sqrt[3]{x^2}$ is defined for all values of x and is even; its graph is shown in Fig. 19. Consequently, its inverse $y = x^{\frac{3}{2}} = \sqrt{x^3}$ has two single-valued branches defined in the interval $[0, \infty)$; their graphs are given in dotted line in Fig. 19. The function $y = x^{\frac{5}{3}}$ and its inverse $y = x^{\frac{3}{5}}$ are odd and defined on the entire x -axis. Their graphs are similar to those in Fig. 15. It is clear that for other positive rational exponents the graphs of the corresponding power functions are of one of these types.

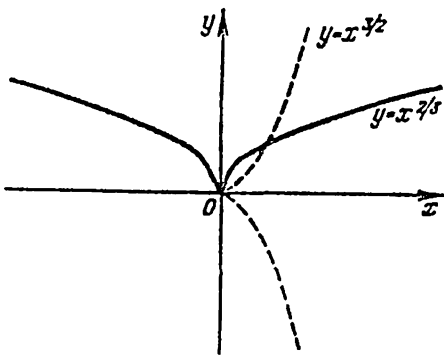


Fig. 19

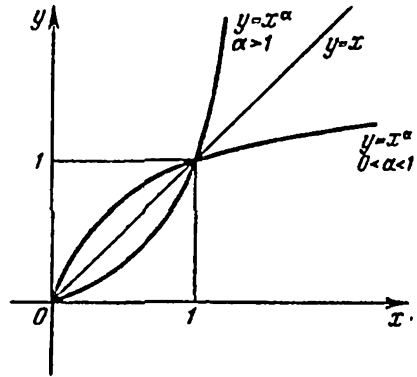


Fig. 20

For irrational (and positive) values of the exponent α the power function is only considered for $x \geq 0$. (Its definition and investigation for $x < 0$ involves complex numbers and functions, and we do not consider this question in our course.)

Fig. 20 shows schematically the graphs of the functions $y = x^\alpha$ for $\alpha < 1$ and $\alpha > 1$ in the first quadrant. These functions are increasing in the interval $[0, \infty)$; their graphs pass through the points $(0, 0)$ and $(1, 1)$, and the straight line $y = x$ separates them into two classes: for $\alpha < 1$ they are convex upward and for $\alpha > 1$ convex downward.

The curve $y = x^\alpha$ with $\alpha > 0$ is called a *parabola of order α* . Thus, the graph of $y = x^2$ (which is an ordinary parabola) is a parabola of the second order, and $y = x^3$ describes a parabola of the third order. The curve $y = x^{\frac{3}{2}}$ (or $y^2 = x^3$) which is frequently encountered is a parabola of order $\frac{3}{2}$ and is referred to as the *semi-cubical parabola* (Fig. 19).

For $\alpha < 0$ the graphs of the functions $y = x^\alpha$ are of an essentially different type. Let $\alpha = -\beta$ where $\beta > 0$. In Fig. 21 we see

the graphs of the functions $y = \frac{1}{x^\beta}$ (for various $\beta > 0$) in the first quadrant. All these functions are decreasing in the interval $(0, \infty)$. When x increases unlimitedly the variable y decreases and approaches zero indefinitely, and, conversely, when y increases unlimitedly the value of x decreases and approaches zero indefinitely. It follows that for any $\beta > 0$ the positive halves of the axes Ox

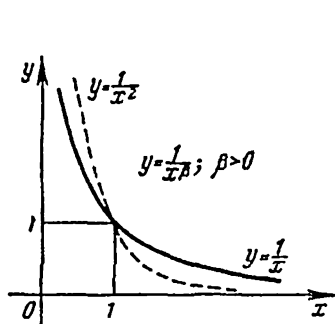


Fig. 21

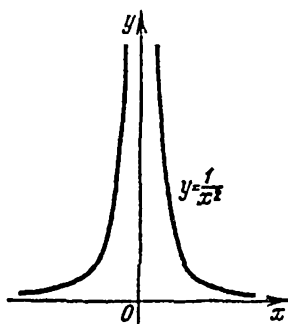


Fig. 22

and Oy are asymptotes of the line $y = \frac{1}{x^\beta}$. As before, all the curves pass through the point $(1, 1)$.

In Fig. 22 the graph of the function $y = \frac{1}{x^2}$ is shown in the whole domain of definition. Let the reader construct the graphs of the functions $y = \frac{1}{x^\beta}$ for various *rational* exponents β .

The line $y = \frac{1}{x^\beta}$, $\beta > 0$, is called a *hyperbola of order β* ; for $\beta = 1$ it is an ordinary equilateral hyperbola.

In conclusion we note that in physics and engineering the lines $y = x^\alpha$ are referred to as *polytropic curves*.

20. Exponential and Logarithmic Functions.

1. Exponential function. The exponential function

$$y = a^x$$

is considered only for $a > 0$ and $a \neq 1$.¹

This function is defined on the whole x -axis and is everywhere positive, i.e. $a^x > 0$ for any x . This means that when x is a fraction

¹ For $a < 0$ the domain of definition of the function $y = a^x$ consists solely of rational numbers $x = \frac{p}{q}$ with odd denominators q ; for $a = 1$ the function is equal to 1 for all values of x .

we confine ourselves to the *arithmetic* root. Therefore the graph of any exponential function lies above Ox . Since $a^0 = 1$, it passes through the point $(0, 1)$. The behaviour of an exponential function depends on whether $a > 1$ or $a < 1$. If $a > 1$ an increase of the exponent x results in an increase of y , and if the argument increases indefinitely the same is with the function. If $a < 1$ then, conversely, as the argument increases, the function decreases and approaches indefinitely (asymptotically) zero.

The graphs of the exponential functions with bases a and $\frac{1}{a}$ are symmetric with respect to Oy . For the function $y = \left(\frac{1}{a}\right)^x$ can be written as $y = a^{-x}$ whence it follows that for positive x 's it assumes the same values as the function $y = a^x$ for the negative x 's having the same absolute values, and vice versa. This exactly means that the graphs of the functions

$$y = a^x$$

and

$$y = \left(\frac{1}{a}\right)^x = a^{-x}$$

are symmetric relative to the axis of ordinates.

In Fig. 23 we see the graphs of two exponential functions for $a > 1$ and $a < 1$. The x -axis serves as the asymptote for the line $y = a^x$.

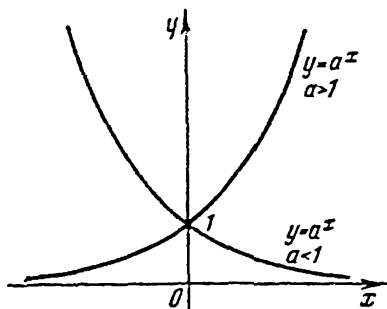


Fig. 23

Exponential functions are encountered in various problems. Most frequently we deal with the exponential function whose base is a special number e playing a very important role in mathematics (see Sec. 31); its approximate value is 2.718. The exponential function $y = e^x$ is also designated as $y = \exp x$. There are many

important relationships in natural sciences that are expressed by formulas of the type $y = Ae^{kx}$.

II. Logarithmic function. The logarithmic function

$$y = \log_a x$$

is the inverse of the exponential function $y = a^x$ ($a > 0$, $a \neq 1$). Therefore, for a given a , it is easy to construct its graph by rotating Fig. 23 about the bisector of the first and third quadrants through 180° . The graph of a logarithmic function is called a *logarithmic curve* (see Fig. 24).

The graphs in Fig. 24 make it possible to describe the properties of logarithmic functions. First of all, we see that every logarithmic function is defined throughout the positive x -axis and is not defined for the negative and zero values of the independent variable. All logarithmic curves pass through the point $(1, 0)$ since the logarithm of unity is always equal to zero.

Properties of a logarithmic function are essentially dependent on whether $a > 1$ or $a < 1$. In the former case ($a > 1$) the logarithmic function is *increasing* over the whole interval $(0, \infty)$, assumes negative values in the interval $(0, 1)$ and positive on the interval $(1, \infty)$. In the latter case ($a < 1$) the logarithmic function is *decreasing* in the whole interval $(0, \infty)$, is positive in the interval $(0, 1)$ and negative in the interval $(1, \infty)$.

The graphs of the logarithmic functions to bases a and $\frac{1}{a}$ are symmetric with respect to Ox .

Taking into account that the logarithmic and exponential functions are mutually inverse we can write (see Sec. 18)

$$\log_a a^x = x \text{ and } a^{\log_a x} = x$$

The first of these equalities is valid for any x and the second only for $x > 0$.

Using the last relation we can represent, for $x > 0$, any power function $y = x^\alpha$ with an arbitrary exponent α in the form of a function composed of the logarithmic and exponential functions:

$$y = x^\alpha = (a^{\log_a x})^\alpha = a^{\alpha \log_a x}$$

Let us take logarithmic functions to two different bases a_1 and a_2 : $y_1 = \log_{a_1} x$ and $y_2 = \log_{a_2} x$. Expressing x from the first relation in terms of y_1 and substituting it into the second relation we obtain $y_2 = \log_{a_2} a_1^{y_1} = y_1 \log_{a_2} a_1$ whence

$$\log_{a_2} x = \frac{\log_{a_1} x}{\log_{a_2} a_1}$$

Putting $x = a_2$ we receive the relation $\log_{a_2} a_2 = \frac{1}{\log_{a_2} a_1}$.

In order to pass in the logarithmic function from a base a_2 to another base a_1 we should multiply the logarithms to the base a_2 by the constant factor $\frac{1}{\log_{a_2} a_1}$. Consequently, the *logarithms of the numbers to different bases are proportional*, that is one logarithmic

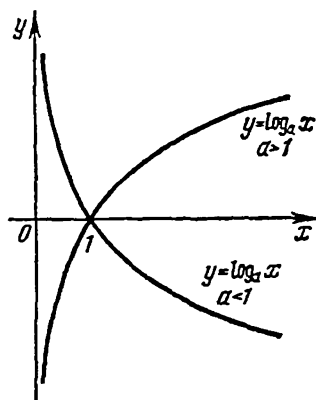


Fig. 24

curve is transformed into another by increasing or decreasing all its ordinates the same number of times.

The logarithms to the base 10 are denoted as $\log x$ and are called *common* (or *decimal*) logarithms, and the logarithms to the base e are termed *natural* logarithms¹ (see Sec. 37, III) and denoted as $\ln x$.

§ 5. Trigonometric, Inverse Trigonometric, Hyperbolic and Inverse Hyperbolic Functions

21. Trigonometric Functions. Harmonic Vibrations.

1. Trigonometric functions. In mathematical analysis the argument of the trigonometric functions

$$\begin{aligned} y &= \sin x, & y &= \cos x, & y &= \tan x \\ y &= \sec x, & y &= \csc x, & y &= \cot x \end{aligned}$$

is always considered as an arc or angle measured in radians. For instance, the value of the function $y = \sin x$ corresponding to $x = x_0$ is equal to the sine of the angle of x_0 radians.

Remember that the *radian measure* of an angle is the ratio of the length of the arc it subtends to the radius of the circle in which it is the central angle (a constant ratio for all such circles). In this very sense such expressions as $x + \sin x$, $x \cos x$ and the like are understood. We suggest that the reader compute the values of these expressions for $x = 0.5$, $x = 1$ and $x = 2$ using trigonometric tables.

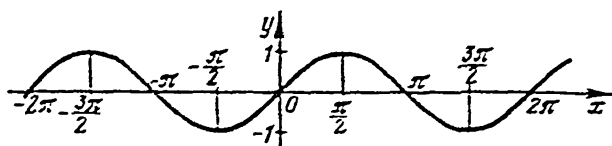


Fig. 25

The six trigonometric functions $y = \sin x$, $y = \cos x$, $y = \tan x$, $y = \sec x$, $y = \csc x$ and $y = \cot x$ are connected by the five well-known algebraic relations (derived from the definitions of these functions) which make it possible to determine from a value of one of them the corresponding values of the others.

The trigonometric functions are periodic: the functions $\sin x$ and $\cos x$ (and therefore $\csc x$ and $\sec x$ as well) with period 2π and the function $\tan x$ (and therefore $\cot x$) with period π .

Let us consider the well-known graphs of trigonometric functions. We begin with the graph of the function $y = \sin x$ (see Fig. 25).

¹ Natural logarithms are also referred to as *Napierian* after J. Napier (1550-1617), a Scottish mathematician.

In the interval $\left[0, \frac{\pi}{2}\right]$ it increases from zero to unity and then in the interval $\left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$ decreases passing through zero attained at the point $x=\pi$ and reaching its least value -1 at $\frac{3\pi}{2}$. Finally, in the interval $\left[\frac{3\pi}{2}, 2\pi\right]$ it again increases from -1 to zero. The period of the function $y=\sin x$ being 2π , the whole graph (called the *sinusoid*) is obtained by shifting its part corresponding to the interval $[0, 2\pi]$ to the right and to the left by $2\pi, 4\pi, 6\pi, \dots$. The function $y=\sin x$ is odd, which is clearly seen in Fig. 25 since the graph is symmetric about the origin.

For the cosine we have

$$\cos x = \sin\left(x + \frac{\pi}{2}\right)$$

and hence, according to Sec. 14, the graph of the function $y=\cos x$ is obtained from the sinusoid by shifting the latter $\frac{\pi}{2}$ units of length to the left along the x -axis (see Fig. 26).

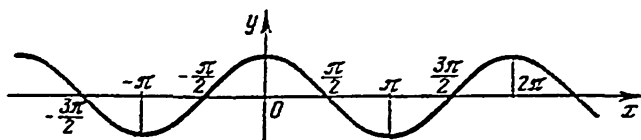


Fig. 26

The function $y=\cos x$ decreases on the interval $[0, \pi]$ from 1 to -1 assuming zero value at the point $x=\frac{\pi}{2}$ and then, on the interval $[\pi, 2\pi]$, increases from -1 to 1 passing through zero value at the point $x=\frac{3\pi}{2}$. On the rest of the number scale its graph is obtained by the periodic continuation with period 2π . The function $y=\cos x$ is even, and its graph is symmetric with respect to the axis of ordinates.

The function $y=\tan x$ is periodic with period π , and therefore it is sufficient to analyse the shape of its graph on the interval $[0, \pi]$. It is not defined at the point $x=\frac{\pi}{2}$; when x increases and approaches $\frac{\pi}{2}$ from the left the function $y=\tan x$ is positive and increases indefinitely; as x decreases and approaches $\frac{\pi}{2}$ from the right the function $y=\tan x$ is negative and its absolute value increases indefinitely. Because of the periodicity of the function

$y = \tan x$ with period π the shape of the graph is the same in the vicinity of every point $x = (2k+1)\frac{\pi}{2}$ where k is an arbitrary integer. The graph of the function $y = \tan x$ (called the *tangent curve*) on the entire axis Ox is obtained by the periodic continuation (with period π) of its part corresponding to the interval $[0, \pi]$ (see Fig. 27). On every interval where the function $y = \tan x$ is

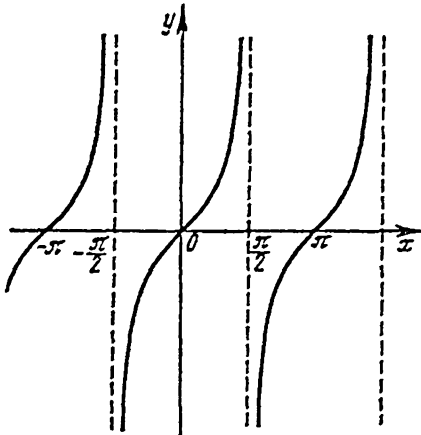


Fig. 27

defined it *increases*. It is an odd function, and therefore its graph is symmetric with respect to the origin.

We suggest that the reader draw the graphs of the remaining trigonometric functions $\csc x$, $\sec x$ and $\cot x$ and describe their peculiarities.

II. Harmonic vibrations. Trigonometric functions have many important applications in mathematics, natural sciences and engineering. They are mainly involved in problems dealing with *periodic processes* whose course is repeated in one and

the same succession and form over certain intervals of the argument (which most often is time t).

The simplest processes of that kind are the so-called *harmonic vibrations* in which the coordinate s of a vibrating point depends on time t according to the *harmonic law*

$$s = A \sin(\omega t + \varphi_0)$$

This function describes a *sinusoidal* functional relationship in which the constant number A is called the *amplitude of vibrations*.¹ The number A is equal to the greatest value of $|s|$ attained in the process. The argument $\omega t + \varphi_0$ substituted into the sine is referred to as the *phase of vibrations*, and the number φ_0 which is equal to the value of the phase for $t=0$ is the *initial phase*. Finally, $\frac{\omega}{2\pi}$ is termed the *frequency of vibrations*.

The last term is accounted for by the relationship between the quantity ω and the period T of the function in question (T is the

¹ For definiteness, we shall suppose that A is positive; if the factor A in the expression $A \sin(\omega t + \varphi_0)$ is negative we can write

$$s = A' \sin(\omega t + \varphi'_0)$$

where $A' = -A > 0$ and $\varphi'_0 = \varphi_0 + \pi$.

period of vibrations). The function $s = A \sin(\omega t + \varphi_0)$ is periodic, its period T being equal to $\frac{2\pi}{\omega}$ since the phase gains an increment of 2π if the independent variable t is given the increment $\frac{2\pi}{\omega}$.

Hence, $\frac{\omega}{2\pi} = \frac{1}{T}$, and therefore $\frac{\omega}{2\pi}$ is equal to the number of periods contained in unit time. Thus, it indicates how many times the periodic process in question repeats its shape during a unit time interval, that is, it characterizes the *frequency* of the process.

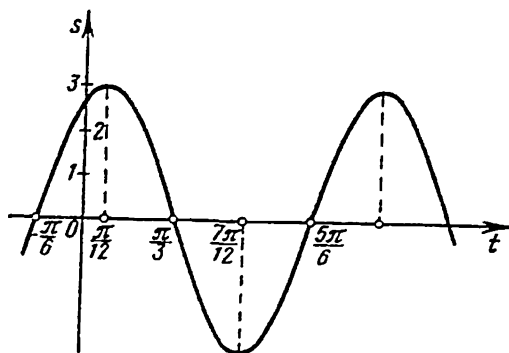


Fig. 28

Accordingly, the number ω shows how many times the repetition occurs during an interval of 2π time units; it is called the *circular frequency* or, briefly, the *frequency*.

To construct the graph of the function

$$s = A \sin(\omega t + \varphi_0)$$

it is convenient to write it down in the form $s = A \sin\left(t + \frac{\varphi_0}{\omega}\right)$ and to apply the techniques described in Sec. 14.

We suggest that the reader construct the graph of the function

$$s = 3 \sin\left(2t + \frac{\pi}{3}\right)$$

(its sketch is shown in Fig. 28).

A vibration described by the equation

$$s = A \sin(\omega t + \varphi_0)$$

is called a *simple harmonic motion (vibration)*, its graph being referred to as a *harmonic*.

In applications we often deal with vibrations of a more complicated form representable as a sum of simple harmonics:

$$s = A_1 \sin(\omega_1 t + \varphi_1) + A_2 \sin(\omega_2 t + \varphi_2) + \dots + A_n \sin(\omega_n t + \varphi_n)$$

where $A_1, A_2, \dots, A_n$ are constants. A vibration of this type is called a *superposition* of simple harmonic motions (a *compound harmonic*). The graph of such a vibration may be of an extremely complicated shape. In Fig. 29 we see the graph of the function

$$s = 2\sin 2t + \sin\left(3t + \frac{\pi}{2}\right)$$

given in continuous line and the graphs of the constituent functions $2\sin 2t$ and $\sin\left(3t + \frac{\pi}{2}\right)$ in dotted line. This is a periodic function

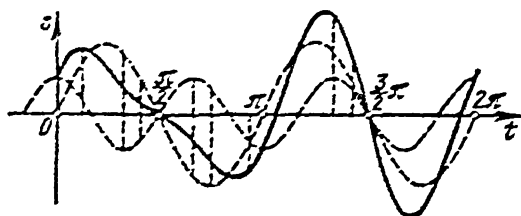


Fig. 29

with period 2π represented as superposition of two simple harmonics

$$s = 2\sin 2t \text{ and } s = \sin\left(3t + \frac{\pi}{2}\right)$$

22. Inverse Trigonometric Functions. Inverse trigonometric functions are investigated with the aid of the results of Sec. 18. Let us begin with the inverse of the function $y = \sin x$. The domain of definition of $\sin x$ (which is the whole number scale) is divided into an infinite number of intervals of monotonicity:

$$\dots, \left[-\frac{3\pi}{2}, -\frac{\pi}{2}\right], \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \left[\frac{\pi}{2}, \frac{3\pi}{2}\right], \dots$$

As a principal interval let us take $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. The branch of the inverse of the function $y = \sin x$ corresponding to this interval is denoted as $y = \arcsin x$ (read "arc sine of x "). The range of the function $y = \sin x$ being the interval $[-1, 1]$, the *domain of definition* of the function $y = \arcsin x$ coincides with that interval. The *range* of the function $y = \arcsin x$ is the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$:

$$-\frac{\pi}{2} \leq \arcsin x \leq \frac{\pi}{2}$$

The value of the function $\arcsin x$ is equal to the magnitude of the angle measured in radians whose sine is equal to the given value of the independent variable x , this angle being chosen among all the angles satisfying this condition so that it lies between

$-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Thus, the equality $y = \arcsin x$ is equivalent to the two equalities

$$\sin y = \sin(\arcsin x) = x, \quad -\frac{\pi}{2} \leq y \leq \frac{\pi}{2}$$

Constructing the graph of $y = \arcsin x$ according to the general rule for determining the graph of an inverse function, that is rotating Fig. 25 through 180° about the bisector of the first and third quadrants, we obtain the curve given in heavy line in Fig. 30. It is readily seen that the function $y = \arcsin x$ is *increasing* and *odd* ($\arcsin(-x) = -\arcsin x$).

Now, we construct on every interval of monotonicity of $y = \sin x$ the corresponding inverse function and thus obtain an infinitude of single-valued branches each of which is defined in the same interval $[-1, 1]$. The graphs of two of them corresponding to the intervals $[-\frac{3\pi}{2}, -\frac{\pi}{2}]$ and $[\frac{\pi}{2}, \frac{3\pi}{2}]$ are shown in Fig. 30. Let the reader check that the first of these functions is equal to $-\pi - \arcsin x$ and the second to $\pi - \arcsin x$. Also write the expressions for two more branches corresponding to the neighbouring intervals. The collection of all the single-valued branches of the inverse of $y = \sin x$ is denoted by

$$y = \text{Arc sin } x$$

According to the terminology introduced in previous sections we can say that this is an *infinite-valued* function since it splits into an infinite number of single-valued branches. If the graphs of all of them are drawn this results in a sinusoid similar to the one in Fig. 25 but "oscillating" about the axis of ordinates (see Fig. 30).

As a rule we deal with the branch $y = \arcsin x$ called the *principal value* of the function $\text{Arc sin } x$.

The inverse of the function $y = \cos x$ is investigated in a completely similar manner. The intervals of monotonicity of $y = \cos x$ are

$$\dots, [-2\pi, -\pi], [-\pi, 0], [0, \pi], [\pi, 2\pi], \dots$$

The inverse of the function $y = \cos x$ corresponding to the interval $[0, \pi]$ is denoted as $y = \arccos x$ (read "arc cosine of x "). Its graph is given in heavy line in Fig. 31. This function is defined in the interval $[-1, 1]$ and takes on the values lying between 0 and π :

$$0 \leq \arccos x \leq \pi$$

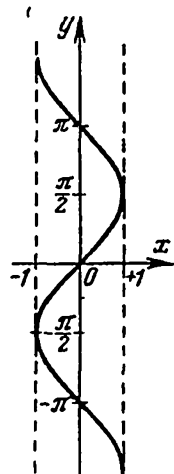


Fig. 30

Consequently, the equality $y = \arccos x$ is equivalent to the two relations

$$\cos y = \cos(\arccos x) = x \text{ and } 0 \leq y \leq \pi$$

The function $y = \arccos x$ is *decreasing* since $\cos x$ decreases in the interval $[0, \pi]$. When computing the values of $y = \arccos x$ it is advisable to use the identity

$$\arccos(-x) = \pi - \arccos x$$

which is directly implied by the reduction formula $\cos(\pi - \alpha) = -\cos \alpha$.

The collection of all single-valued branches of the inverse of $y = \cos x$ (their number is again infinite) is denoted as $y = \text{Arccos } x$;

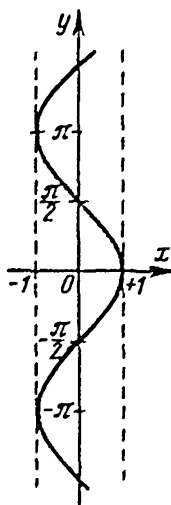


Fig. 31

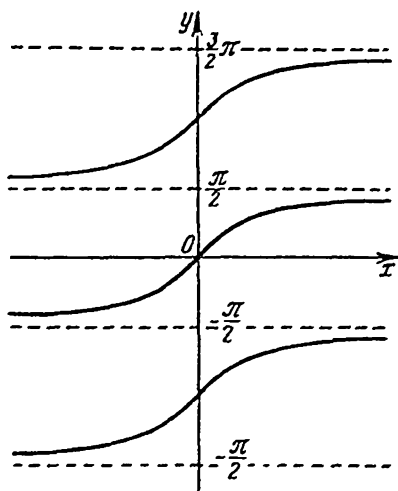


Fig. 32.

the graphs of these branches are shown in Fig. 31, the branch $y = \arccos x$ being called the *principal value* of $y = \text{Arccos } x$.

We also give the following formulas:

$$\sin(\arccos x) = +\sqrt{1-x^2}$$

$$\cos(\arcsin x) = +\sqrt{1-x^2}$$

$$\arcsin x + \arccos x = \frac{\pi}{2}$$

(let the reader derive them).

Now let us consider the inverse of the function $y = \tan x$. The function increases on the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$, and consequently it possesses a single-valued inverse function which is denoted as $y = \arctan x$ ("arc tangent of x "). Its graph is given in heavy line in Fig. 32. The properties of the function $\tan x$ imply that the

function $y = \arctan x$ is defined on the *entire* number scale and is *increasing* and *odd*. Its range is the interval $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$:

$$-\frac{\pi}{2} < \arctan x < \frac{\pi}{2}$$

The collection of all single-valued branches of the inverse of $y = \tan x$ is designated as $y = \text{Arctan } x$. The graphs of the branches different from $y = \arctan x$ (which is termed the *principal value* of $\text{Arctan } x$) are obtained from the graph of $y = \arctan x$ by parallel displacements along Oy by $k\pi$, $k = \pm 1, \pm 2, \dots$ (Fig. 32).

Let the reader also analyse the graph of the inverse of the function $y = \cot x$.

In what follows, unless otherwise stated, "inverse trigonometric functions" means the principal branches of the functions.

Trigonometric functions are also called *circular functions* and their inverses are spoken of as *inverse circular functions*.

23. Hyperbolic and Inverse Hyperbolic Functions. The hyperbolic functions are not included into the class of basic elementary functions but we discuss them here since they are important for applications and admit of a simple investigation. They are encountered in various applied sciences (in particular, in electrical engineering, strength of materials and the like).

1. Definition. The functions

$$\cosh x = \frac{e^x + e^{-x}}{2}$$

$$\sinh x = \frac{e^x - e^{-x}}{2}$$

and

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

where $e = 2.718 \dots$ (see Sec. 31) are, respectively, called the *hyperbolic cosine*, the *hyperbolic sine* and the *hyperbolic tangent*.

These functions are defined for all values of x . They are connected by a number of algebraic relations similar to those connecting the corresponding trigonometric functions (these relations account for the words "cosine", "sine" and "tangent" entering into the names of the functions). In particular, there hold the formulas

$$\cosh^2 x - \sinh^2 x = 1, \quad \cosh 2x = \cosh^2 x + \sinh^2 x,$$

$$\sinh 2x = 2 \sinh x \cosh x,$$

$$\cosh^2 x = \frac{1}{1 - \tanh^2 x}, \quad \sinh^2 x = \frac{\tanh^2 x}{1 - \tanh^2 x}$$

which are readily checked.

Here we give without proof the following simple rule: if in any identity connecting the trigonometric functions $\cos x$, $\sin x$, and $\tan x$ the expressions $\cos x$, $\sin x$ and $\tan x$ are, respectively, replaced by $\cosh x$, $i \sinh x$ and $i \tanh x$ where $i^2 = -1$, an identity for the hyperbolic functions is obtained.

The graphs of the functions $\cosh x$ and $\sinh x$ are easily constructed since we know the graphs of the functions e^x and e^{-x} (see Sec. 20); to this end we apply the method indicated at the end of Sec. 14. The construction of the graph of the function

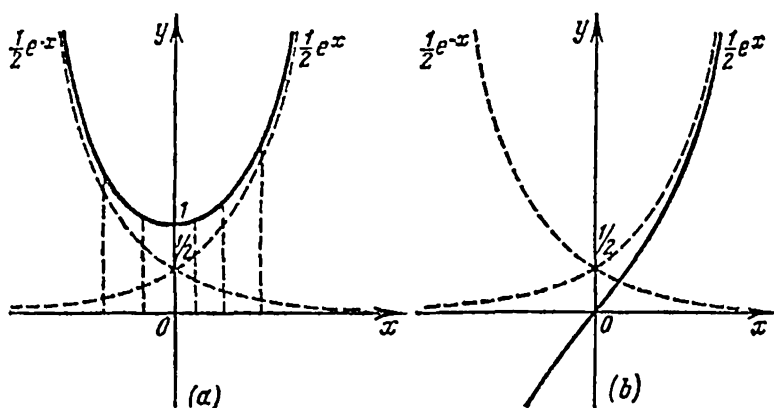


Fig. 33

$y = \cosh x$ is demonstrated in Fig. 33a. This function is *even* and *positive*. The function $y = \sinh x$ is *odd*; it is positive for $x > 0$, negative for $x < 0$ and assumes zero value at the point $x = 0$. Its graph is shown in Fig. 33b.

From the graphs and also from the formulas defining the hyperbolic functions it is seen that for large positive values of x we can approximately write

$$\cosh x \approx \frac{e^x}{2} \text{ and } \sinh x \approx \frac{e^x}{2}$$

since the term $\frac{e^{-x}}{2}$ is negligibly small for large $x > 0$. Even for $x > 4$ the absolute error of these approximations is less than 0.01, the relative error being less than 0.04%.

It clearly follows that for large positive values of x the function $\tanh x$ approaches unity remaining less than unity. The function $y = \tanh x$ being *odd* and equal to zero at the point $x = 0$, its graph can be easily constructed schematically (see Fig. 34). The shape of this graph resembles that of the graph of $\arctan x$ (see Fig. 32), but it should be taken into account that its asymp-

totes are the straight lines $y = \pm 1$ and not the lines $y = \pm \frac{\pi}{2}$ (which are the asymptotes of $\arctan x$).

II. Inverse hyperbolic functions. The inverse functions of the corresponding hyperbolic functions are denoted as

$$y = \cosh^{-1} x, \quad y = \sinh^{-1} x \text{ and } y = \tanh^{-1} x$$

and termed *arc-hyperbolic functions*.

The function $\sinh x$ is defined and increases in the interval $(-\infty, \infty)$, its range coinciding with Oy (Fig. 33b). Therefore the function $y = \sinh^{-1} x$ is defined throughout Ox and increases. It can be expressed explicitly by means of the logarithmic function. To this end we rewrite the relation $y = \sinh^{-1} x$ in the form

$$x = \sinh y = \frac{e^y - e^{-y}}{2}$$

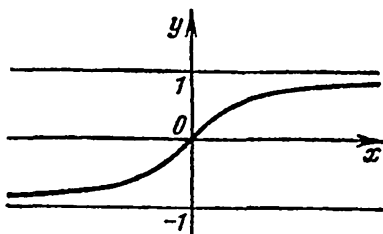


Fig. 34

and solve the resultant equation with respect to e^y and then with respect to y . After some simple transformations this leads to $e^{2y} - 2xe^y - 1 = 0$ whence $e^y = x + \sqrt{x^2 + 1}$. The second root $x - \sqrt{x^2 + 1}$ of the quadratic equation is discarded since it is negative for all values of x and therefore cannot be equal to the positive quantity e^y . Using natural logarithms (Sec. 20, II) we finally obtain

$$y = \sinh^{-1} x = \ln(x + \sqrt{x^2 + 1}) \quad (*)$$

The function $\cosh x$ decreases throughout the interval $(-\infty, 0]$ and increases throughout $[0, \infty)$. Its range is the interval $[1, \infty)$. Consequently, the function $y = \cosh^{-1} x$ splits into two single-valued branches defined for $x \geq 1$. Carrying out computations similar to the derivation of formula (*) we obtain the two expressions

$$(\cosh^{-1} x)_1 = \ln(x - \sqrt{x^2 - 1}) \text{ and } (\cosh^{-1} x)_2 = \ln(x + \sqrt{x^2 - 1})$$

The first branch is the inverse of the function $y = \cosh x$ in the interval $(-\infty, 0]$ while the second corresponds to the interval $[0, \infty)$. It is readily seen that

$$x - \sqrt{x^2 - 1} = \frac{1}{x + \sqrt{x^2 - 1}}$$

that is, $\ln(x - \sqrt{x^2 - 1}) = -\ln(x + \sqrt{x^2 - 1})$. This is of course a mere consequence of the fact that $y = \cosh x$ is an even function.

The function $y = \tanh^{-1} x$ is defined in the interval $(-1, 1)$, which is the range of the function $\tanh x$. It is easy to verify th

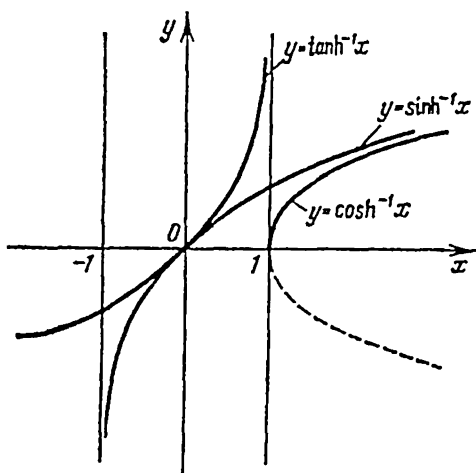


Fig. 35

$$\tanh^{-1} x = \frac{1}{2} \ln \frac{1+x}{1-x}$$

The graphs of the inverse hyperbolic functions are shown in Fig. 35.

QUESTIONS

1. What numbers form the set of real numbers?
2. How is the number system constructed?
3. What do we call an interval?
4. State the definition of neighbourhood of a point.
5. What is the absolute value of a number?
6. State the basic propositions of absolute values.
7. What is the absolute error? What is the limiting absolute error?
8. What is the relative error? What is the limiting relative error?
9. State the definition of a function of one independent variable. What do we call the domain of definition of a function and the range of a function?
10. Give examples of functions of an integral argument.
11. State the definition of the graph of a function in Cartesian coordinates.
12. What does "to specify a function" mean? Describe the tabular, analytical and graphical methods of representing functions. What advantages and demerits does each of these methods possess?
13. Describe various types of notation of functions.
14. What do we call a particular value of a function? How is it denoted?
15. Enumerate the basic elementary functions.
16. What is a composite function? Give examples of composite functions.
17. Describe the class of elementary functions.
18. State the definitions of rational, algebraic and transcendental functions.
19. What is an implicit function? Give examples of implicit functions.

20. What do we call a zero of a function?
21. State the definitions of even and odd functions.
22. What do we call a periodic function? What is a period of a function?
23. State the definition of a function increasing (decreasing) in an interval.
24. What is an interval of monotonicity of a function?
25. State the definitions of the greatest and the least values of a function on an interval.
26. Given the graph of a function $y=f(x)$, describe the construction of the graph of the function $y=a_1f(a_2x+a_3)+a_4$ where a_1 , a_2 , a_3 and a_4 are constants.
27. State the definition of a linear function. Draw its graph and prove its basic property.
28. What is a quadratic function? Draw its graph and describe its behaviour.
29. State the definition of a linear-fractional function. Draw its graph and describe its behaviour.
30. What do we call the inverse of a given function? How is the graph of the inverse function constructed in Cartesian coordinates when the graph of the direct function is given?
31. How are single-valued branches of the inverse of a non-monotone function defined?
32. Draw the graphs of power functions for various exponents and describe the properties of these functions.
33. Draw the graphs of exponential functions to various bases and describe the properties of these functions.
34. Draw the graphs of logarithmic functions to various bases and describe the properties of these functions.
35. Draw the graphs of the trigonometric functions and describe their properties.
36. What is a harmonic vibration? What do we call the amplitude, the phase, the initial phase, the period, the frequency and the circular frequency of a simple harmonic motion?
37. How do we construct a simple harmonic?
38. Draw the graphs of the inverse trigonometric functions and describe the properties of these functions.
39. How are the hyperbolic functions defined? Draw the graphs of these functions and describe their properties.
40. Describe the properties of the inverse hyperbolic functions and draw their graphs.

Chapter II

LIMIT. CONTINUITY

§ 1. Limit. Infinitely Large Magnitudes

24. Limit of a Function. In this section we consider a function $y=f(x)$ of a continuous argument x (see Sec. 7). Let us introduce the notion of a *limit of a function* playing a fundamental role in mathematical analysis.

Suppose that the independent variable x approaches indefinitely a number x_0 . This means that x is made to assume the values which become arbitrarily close to x_0 but are not equal to x_0 . To describe such a situation we say that x tends to x_0 and write $x \rightarrow x_0$. It may turn out that in this process the values of $f(x)$ corresponding to $x \rightarrow x_0$ approach indefinitely a number A . Then we say that the number A is the *limit* of the function $y=f(x)$ as $x \rightarrow x_0$ or that the function $y=f(x)$ tends to the number A as $x \rightarrow x_0$.

Definition. The number A is said to be the *limit of the function* $y=f(x)$ as $x \rightarrow x_0$ if for all the values of x lying sufficiently close to x_0 the corresponding values of the function $f(x)$ are arbitrarily close to the number A .

If A is the limit of the function $f(x)$ as $x \rightarrow x_0$ we write

$$\lim_{x \rightarrow x_0} f(x) = A \text{ or } f(x) \rightarrow A_{x \rightarrow x_0}$$

(the abbreviation "lim" means "limit").

It should be noted that this definition does not require that the function should be defined at the point x_0 itself. It is only necessary that the function should be defined in an arbitrary neighbourhood of this point; at the point x_0 it may not be defined. For instance, the function $y = \frac{\sin x}{x}$ is not defined at the point $x=0$ but, as will be shown in Sec. 30, it tends to 1 as x tends to zero.

The determination of the limit of a function defined in a neighbourhood of a point x_0 , except possibly at the point x_0 itself, is one of the most important problems of the theory of limits.

The fact that a number A is the limit of a function $f(x)$ as $x \rightarrow x_0$ can be checked with the aid of the following test (which is equivalent to the definition of the limit):

If, given any positive number ε (however small), there is a positive number δ such that for all x different from x_0 and satisfying the inequality

$$|x - x_0| < \delta$$

(that is, for all $x \neq x_0$ belonging to the δ -neighbourhood of the point x_0) *there holds the inequality*

$$|f(x) - A| < \varepsilon$$

the number A is the limit of the function $f(x)$ as $x \rightarrow x_0$.

As will be shown, in the concrete problems we shall be concerned with there is no practical need in the application of this test which involves lengthy and complicated calculations, since we shall derive some simple rules for finding limits.

As an example, let us take the function $y = 3x - 1$ and prove that for $x \rightarrow 1$ it has the limit equal to 2. To this end, choose a positive number ε . For the inequality

$$|(3x - 1) - 2| < \varepsilon \quad (*)$$

or the equivalent inequality

$$3|x - 1| < \varepsilon$$

to be fulfilled it is sufficient that the condition

$$|x - 1| < \frac{\varepsilon}{3}$$

should hold.

Consequently, for all x differing from 1 by less than $\frac{\varepsilon}{3}$ the function differs from 2 by less than ε where ε is an arbitrary positive number. Therefore, the function $y = 3x - 1$ does in fact tend to 2 as $x \rightarrow 1$. In this example we can put $\delta = \frac{\varepsilon}{3}$. Here it is not necessary to stipulate the condition $x \neq 1$ since inequality (*) is also fulfilled for $x = 1$.

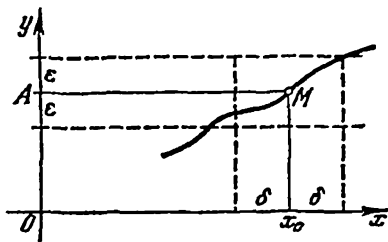


Fig. 36

The existence of the limit, equal to A , of a function $y = f(x)$ for $x \rightarrow x_0$, is illustrated geometrically (see Fig. 36) as follows: we erect the perpendiculars to Ox at the point x_0 and to Oy at the point A , extend them to meet at the point M and set an arbitrary positive number ε ; the above test then implies that

there exists a δ -neighbourhood of the point x_0 such that the part of the graph of the function $y=f(x)$ corresponding to this neighbourhood (with the point x_0 deleted) is contained in the strip bounded by the straight lines $y=A-\epsilon$ and $y=A+\epsilon$.

It is evident that if a function has a limit for $x \rightarrow x_0$ this limit is unique since the values of the function corresponding to the values of x approaching x_0 must become arbitrarily close to a constant and hence cannot be simultaneously close to *more than one* constant number.

A constant magnitude C is interpreted as a constant function $y=C$, and, consequently, according to the above test, it possesses a limit equal to its own value:

$$\lim_{x \rightarrow x_0} C = C$$

For the difference $C-C$ is equal to zero and thus is less than any given positive number.

It is readily seen that for any point x_0 the function $y=x$ has a limit which is equal to x_0 :

$$\lim_{x \rightarrow x_0} x = x_0$$

Indeed, the inequality $|x-x_0| < \epsilon$ where ϵ is an arbitrary positive number holds for all x 's belonging to the ϵ -neighbourhood of the point x_0 , and therefore, in this case, we can simply put $\delta = \epsilon$.

The definition of the limit and the above test do not provide practical means for *finding* limits; later on we shall derive some rules facilitating the determination of limits.

Now let us discuss an example of a function having no limit as $x \rightarrow 0$. Consider the function defined by the condition

$$y = \frac{|x|}{x} \quad \text{for } x \neq 0$$

This function is not defined at the point $x=0$. For the values of x less than zero the function is equal to -1 and for the values of x exceeding zero it is equal to 1 . Hence, there exists no number to which the values of the function become arbitrarily close for *all* the values of x approaching the point 0 .

The function appearing if we add the point $x=0$ to the domain of definition of the function considered in the previous example and put $y=0$ for $x=0$ is called the *signum function* and denoted as $y = \operatorname{sgn} x$:

$$y = \operatorname{sgn} x = \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ 1 & \text{for } x > 0 \end{cases}$$

(the abbreviation "sgn" means "signum function"). Its graph is depicted in Fig. 37.

25. **Limits of Functions for $x \rightarrow \pm \infty$.** Let the independent variable x of a function $y = f(x)$ *increase indefinitely*. This means that x is made to take on the values becoming greater than any given positive number. In such a case we say that x becomes *positively infinite* or that x *approaches (tends to) plus infinity* and write $x \rightarrow +\infty$. If x *decreases indefinitely*, that is, becomes less than any given negative number, we say that x *becomes negatively infinite* or that it *approaches minus infinity* and write $x \rightarrow -\infty$.

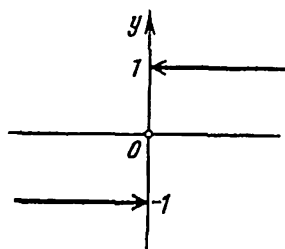


Fig. 37

When studying a function whose argument tends to (plus or minus) infinity we of course suppose that the function is defined for all the values of x under consideration.

It may turn out that for the argument approaching plus (minus) infinity the corresponding values of the function $y = f(x)$ become arbitrarily close to (*tend to*) a number A . Then we say that A is the *limit* of the function $y = f(x)$ for $x \rightarrow +\infty$ ($x \rightarrow -\infty$) or that the function $y = f(x)$ *tends to the number A as $x \rightarrow +\infty$ (as $x \rightarrow -\infty$)*.

Let us begin with the definition of the limit for $x \rightarrow +\infty$.

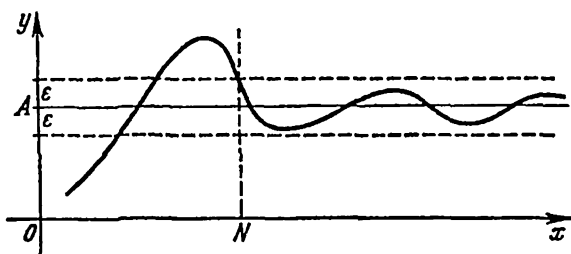


Fig. 38

Definition. A number A is said to be the *limit of the function $y = f(x)$ as $x \rightarrow +\infty$* if for all sufficiently large values of x the corresponding values of the function $f(x)$ become arbitrarily close to the number A .

If A is the limit of a function $f(x)$ as $x \rightarrow +\infty$ we write

$$\lim_{x \rightarrow +\infty} f(x) = A \quad \text{or} \quad f(x) \rightarrow A \quad \text{as } x \rightarrow +\infty$$

The existence of the limit, equal to A , of a function $y = f(x)$ as $x \rightarrow +\infty$ can be demonstrated geometrically (see Fig. 38); let us erect the perpendicular to Oy at the point A and choose an

arbitrary positive number ε ; then there is a number $N > 0$ such that the part of the graph of the function $y = f(x)$ corresponding to the values of x exceeding that number is contained within the strip bounded by the lines

$$y = A - \varepsilon \quad \text{and} \quad y = A + \varepsilon$$

In Fig. 39 we see Graphs I and II of functions which, when approaching their limits, as $x \rightarrow +\infty$, increase and decrease, respectively. Fig. 38 shows the graph of a function which approaches its limit, as $x \rightarrow +\infty$, in an oscillating manner. In the last case the graph intersects the straight line $y = A$, that is, the function assumes the value equal to its limit any number of times. This demonstrates, geometrically, possible ways of approaching a limit for $x \rightarrow \infty$.

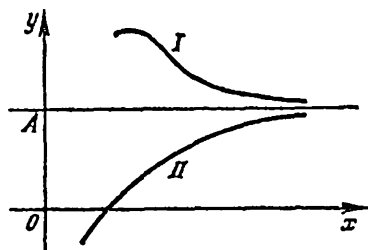


Fig. 39

Bearing in mind the geometrical interpretation of the limit of a function, as $x \rightarrow +\infty$, we can say that

if $\lim_{x \rightarrow +\infty} f(x) = A$, the curve $y = f(x)$ has the straight line $y = A$ as its asymptote for $x \rightarrow +\infty$.

The definition and the geometrical interpretation of the limit of a function as $x \rightarrow -\infty$ are completely analogous to the above, and therefore we leave them to the reader.

It may also occur that a function $f(x)$ tends to one and the same limit A for $x \rightarrow +\infty$ and for $x \rightarrow -\infty$. This means that for all the values of x sufficiently large in their absolute values the corresponding values of $f(x)$ become arbitrarily close to A . In this case we write

$$\lim_{x \rightarrow \infty} f(x) = A \quad \text{or} \quad f(x) \rightarrow A \quad \text{as} \quad x \rightarrow \infty$$

The geometrical significance of this case is that the graph of the function $y = f(x)$ is contained in the strip bounded by the straight lines $y = A - \varepsilon$ and $y = A + \varepsilon$ (where ε is an arbitrarily small positive number) as the point x becomes sufficiently distant from the point $x = 0$ (Fig. 40). In other words, the line $y = A$ is an asymptote of the graph of the function $y = f(x)$ both for $x \rightarrow +\infty$ and for $x \rightarrow -\infty$.

To check that a number A is the limit of the function $f(x)$ as $x \rightarrow +\infty$ (as $x \rightarrow -\infty$) we can use the following test (equivalent to the definition of the limit): if for every positive number ε (however small) there is a positive number N such that for all the values of x satisfying the inequality

$$x > N \quad (x < -N)$$

the inequality

$$|f(x) - A| < \varepsilon \quad (*)$$

holds, the number A is the limit of the function $f(x)$ as $x \rightarrow +\infty$ ($x \rightarrow -\infty$).

If inequality $(*)$ is fulfilled for all x 's satisfying the condition

$$|x| > N$$

then A is the limit of the function both for $x \rightarrow +\infty$ and for $x \rightarrow -\infty$ (i.e. for $x \rightarrow \infty$) since this condition allows x to tend to infinity in an arbitrary fashion.

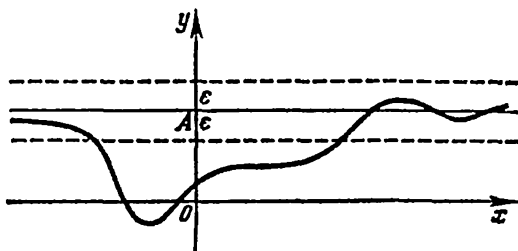


Fig. 40

Examples. (1) The function

$$y = \frac{x+1}{x-1}$$

tends to 1 when x approaches infinity in an arbitrary way, i.e.

$$\lim_{x \rightarrow \infty} \frac{x+1}{x-1} = 1$$

Fig. 14 shows that, as the abscissa of the moving point of the equilateral hyperbola $y = \frac{x+1}{x-1}$ increases indefinitely, this point becomes arbitrarily close to the line $y=1$. Let us apply the above test to verify this conclusion analytically. Choose an arbitrary $\varepsilon > 0$. For the inequality

$$\left| \frac{x+1}{x-1} - 1 \right| = \frac{2}{|x-1|} < \varepsilon$$

or the equivalent inequality

$$|x-1| > \frac{2}{\varepsilon}$$

to hold it is sufficient that

$$|x| - 1 > \frac{2}{\varepsilon}$$

because $|x-1| \geq |x|-1$. Hence, for every x satisfying the inequality

$$|x| > \frac{2}{\varepsilon} + 1$$

the function $\frac{x+1}{x-1}$ differs from 1 by less than ε , which shows that $\lim_{x \rightarrow \infty} \frac{x+1}{x-1} = 1$. Hence we can choose the number N equal to $\frac{2}{\varepsilon} + 1$.

(2) The function $y = \arctan x$ has different limits for $x \rightarrow +\infty$ and $x \rightarrow -\infty$. It is clear from Fig. 32 that

$$\lim_{x \rightarrow +\infty} \arctan x = \frac{\pi}{2} \quad \text{and} \quad \lim_{x \rightarrow -\infty} \arctan x = -\frac{\pi}{2}$$

26. Limit of a Sequence. Let us consider a function of an *integral argument*. Usually such an argument is denoted by the letter n and the values of the function by some other letter supplied with a subscript indicating the value of the integral argument. For instance, if $y = f(n)$ is a function of the integral argument n we write $y_n = f(n)$. Given such a function, we say that the values

$$y_1 = f(1), y_2 = f(2), \dots, y_n = f(n), \dots,$$

assumed by the function form a *sequence*.

Definition. A *sequence* is an infinite set of quantities numbered with the positive integers and ordered as these integers are.

If there is a sequence $y_1, y_2, y_3, \dots$, this assigns, to every natural number n , a value $y_n = f(n)$. For instance, the terms of the geometric progression $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$ are the subsequent values of the function $f(n) = \frac{1}{2^n}$.

It may occur that, as n increases, the values $y_n = f(n)$ become arbitrarily close to a number A . Then we say that the number A is the limit of the function $f(n)$ of the integral argument n or that the sequence $y_1, y_2, \dots, y_n, \dots$ has the limit A , as $n \rightarrow \infty$, and write

$$\lim_{n \rightarrow \infty} f(n) = A \quad \text{or} \quad \lim_{n \rightarrow \infty} y_n = A$$

Definition. A number A is said to be the *limit of the function* $y = f(n)$ of the integral argument n or the *limit of the sequence* $y_1, y_2, \dots, y_n$, if for all sufficiently large integral values of n the corresponding values y_n of the function become arbitrarily close to the number A .

We see that the definition of the limit of a sequence can be regarded as a special case of the definition of the limit of a

function, as its argument becomes positively infinite and assumes only *integral* values. Hence, if A is the limit of a sequence $y_1, y_2, \dots, y_n, \dots$, then, given an arbitrary positive number ε , there is a positive number N such that the inequality

$$|y_n - A| < \varepsilon$$

holds for all $n > N$.

Examples. (1) The limit of the sequence

$$\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots, \frac{n}{n+1}, \dots$$

exists and is equal to 1. Indeed, for the absolute value of the difference between $\frac{n}{n+1}$ and 1, which is equal to

$$\left| \frac{n}{n+1} - 1 \right| = \frac{1}{n+1}$$

to be less than a given positive number ε , it is sufficient that the inequality $n+1 > \frac{1}{\varepsilon}$ should be fulfilled. This is equivalent to the condition $n > \frac{1}{\varepsilon} - 1$, and hence, given ε , there exists $N = \frac{1}{\varepsilon} - 1$ such that for all $n > N$ the absolute value of the above difference is less than ε , which means that 1 is the limit of the sequence in question.

(2) Let us consider the sequence

$$\sin \frac{\pi}{2}, \frac{1}{2} \sin \frac{3\pi}{2}, \frac{1}{3} \sin \frac{5\pi}{2}, \dots, \frac{1}{n} \sin \left[(2n-1) \frac{\pi}{2} \right], \dots$$

The function $y_n = \frac{1}{n} \sin \left[(2n-1) \frac{\pi}{2} \right] = (-1)^{n-1} \frac{1}{n}$ tends to zero as $n \rightarrow \infty$ since $|y_n - 0| = \frac{1}{n} < \varepsilon$ for all values of n satisfying the condition $n > \frac{1}{\varepsilon}$. The difference

$$y_n - 0 = \frac{(-1)^{n+1}}{n}$$

is positive or negative depending on whether n is odd or even. When y_n approaches zero its sign permanently alternates, and the variable tends to the zero limit oscillating about it.

(3) In elementary geometry it is proved that the perimeter of a regular n -gon, inscribed in a given circle, which is a function of the integral argument n has a limit. This limit, by definition, is taken as the circumference of the circle. Similarly, the area of that n -gon possesses a limit which, by definition, is equal to the area of the circle.

(4) Now we give some examples of sequences having no limits.

(a) The sequence

$$\sin \frac{\pi}{2}, \sin \frac{2\pi}{2}, \sin \frac{3\pi}{2}, \dots, \sin \frac{n\pi}{2}, \dots$$

does not possess a limit since, for $n=1, 2, 3, 4$, the variable $y_n = \sin \frac{n\pi}{2}$ consecutively takes on the values 1, 0, -1, 0 which then are repeated indefinitely in the same order as n increases. Therefore the number to which y_n tends, as $n \rightarrow \infty$, does not exist.

(b) The function $y_n = 2n+1$ whose values corresponding to $n=1, 2, 3, \dots$ form the sequence of odd positive integers 1, 3, 5, ... does not tend to any limit as $n \rightarrow \infty$ since y_n increases indefinitely when n becomes positively infinite.

27. **Infinitely Large Magnitudes. Bounded Functions.** For the sake of brevity, we shall only state the definitions for the case when x tends to a *finite* limit x_0 . Let the reader analyse the changes that should be introduced to embrace the case when x becomes infinite.

1. **Infinitely large magnitudes.** Suppose that when $x \rightarrow x_0$ the "limiting behaviour" of a function $y=f(x)$ is such that its absolute value increases unlimitedly. Then we say that the function $f(x)$ is an *infinitely large magnitude* as $x \rightarrow x_0$.

Definition. A function $y=f(x)$ is said to be an *infinitely large magnitude* (or is said to become infinite or to approach infinity) as $x \rightarrow x_0$ if for all the values of x lying sufficiently close to x_0 the moduli of the corresponding values of the function $f(x)$ become greater than any given arbitrarily large positive number.

If a function $f(x)$ approaches infinity as $x \rightarrow x_0$ we write

$$\lim_{x \rightarrow x_0} f(x) = \infty$$

This definition of an infinitely large magnitude indicates that the absolute value of the function $f(x)$ becomes arbitrarily large as the argument x enters a sufficiently small neighbourhood of the point x_0 .

The relationship $\lim_{x \rightarrow x_0} f(x) = \infty$ can be checked by the following test (equivalent to the definition): *if, given any positive number M , however large, there is a positive number δ such that for all the values of x different from x_0 and satisfying the inequality*

$$|x - x_0| < \delta$$

the condition

$$|f(x)| > M$$

is fulfilled, the function $f(x)$ is an infinitely large magnitude as $x \rightarrow x_0$.

In the case $x \rightarrow \infty$ the situation is the same if for every number M , however large, there exists a positive number N such that for all x 's satisfying the condition $|x| > N$ the inequality

$$|f(x)| > M$$

holds.

The fact that a function $y=f(x)$ approaches infinity as $x \rightarrow x_0$ can be illustrated geometrically (see Fig. 41a): if we choose arbitrarily a positive number M then there is a δ -neighbourhood of the point $x=x_0$ such that the part of the graph of the function $y=f(x)$ corresponding to this δ -neighbourhood (with the point x_0 deleted) lies outside the strip bounded by the lines $y=-M$ and $y=M$.

The case $x \rightarrow +\infty$ is demonstrated in Fig. 41b.

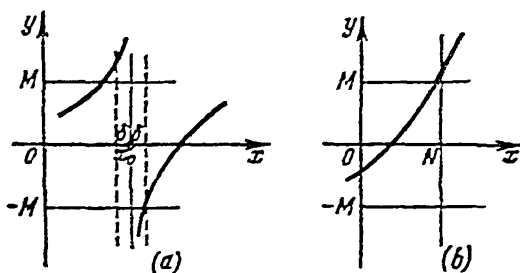


Fig. 41

A function $y=f(x)$ which is an infinitely large magnitude as $x \rightarrow x_0$ does not have a limit in the ordinary sense. To extend the ordinary terminology and to characterize the behaviour of the function whose absolute value $|f(x)|$ increases indefinitely we say that the function $y=f(x)$ tends to infinity or has an infinite limit.

Suppose that a function $y=f(x)$ tending to infinity as $x \rightarrow x_0$ only assumed positive (negative) values in a neighbourhood of the point x_0 . Then we say that the function $f(x)$ becomes positively (negatively) infinite or approaches plus (minus) infinity as $x \rightarrow x_0$. In these cases we write, respectively,

$$\lim_{x \rightarrow x_0} f(x) = +\infty \quad \text{and} \quad \lim_{x \rightarrow x_0} f(x) = -\infty$$

It should be borne in mind that these statements are understood conditionally: infinity is not a number, and therefore it is senseless to speak about ordinary operations on the symbol ∞ .

Also, a concrete constant number, however large, must not be confused with an infinitely large magnitude (which is a variable).

Examples. (1) The function $y = \frac{1}{x}$ approaches infinity as $x \rightarrow 0$. Indeed, for the inequality

$$\frac{1}{|x|} > M$$

(where M is a given positive number) to be valid it is sufficient that the condition

$$|x| < \frac{1}{M}$$

should hold, which means that the values $x \neq 0$ should belong to the $\frac{1}{M}$ -neighbourhood of the point $x=0$ (hence, in this example we can put $\delta = \frac{1}{M}$). Thus,

$$\lim_{x \rightarrow 0} \frac{1}{x} = \infty$$

If x approaches zero and assumes only negative (positive) values the function $y = \frac{1}{x}$ tends to $-\infty$ ($+\infty$):

$$\lim_{\substack{x \rightarrow 0 \\ x < 0}} \frac{1}{x} = -\infty \quad \text{and} \quad \lim_{\substack{x \rightarrow 0 \\ x > 0}} \frac{1}{x} = +\infty$$

This limiting behaviour of the function $y = \frac{1}{x}$ for $x \rightarrow +\infty$ is clearly seen from its graph which is the well-known equilateral hyperbola.

(2) The function $y = \frac{x+1}{x-1}$ is an infinitely large magnitude as $x \rightarrow 1$; it tends to $-\infty$ if x remains less than 1 and to $+\infty$ if x remains greater than 1 (see Fig. 14).

(3) The function $y = \frac{1}{x^2}$ has the limit $+\infty$ as $x \rightarrow 0$ (see Fig. 22).

(4) As can be readily verified, the function $y = a^x$ becomes positively infinite for $x \rightarrow +\infty$ if $a > 1$: $\lim_{x \rightarrow +\infty} a^x = +\infty$. If $a < 1$ we have $\lim_{x \rightarrow -\infty} a^x = +\infty$ (see Fig. 23).

In what follows, when speaking about a limit, we shall mean that the limit is finite unless it is stipulated that it is infinite.

II. Bounded functions.

Definition. A function $y = f(x)$ is said to be *bounded* in a given interval if there is a positive number M such that for all the values of x belonging to this interval the inequality $|f(x)| \leq M$ holds. If otherwise, the function is said to be *unbounded*.

The graph of a bounded function (considered within the given interval) is completely contained between the straight lines $y = -M$ and $y = M$. It is sometimes more convenient to say that a function is bounded if its values are contained between two numbers A and B :

$$A \leq f(x) \leq B$$

(this definition is of course equivalent to the above). Then the graph of the function $y = f(x)$ lies between the lines $y = A$ and $y = B$.

For instance, the functions $y = \sin x$ and $y = \cos x$ are *bounded* throughout Ox while the function $y = a^x$ when considered on the whole x -axis, is *unbounded*. It should be stressed that when speaking of the boundedness of a function it is necessary to indicate the interval on which the function is considered. For example, the function $y = \frac{1}{x}$ is bounded in the interval $(1, \infty)$ and unbounded in the interval $(0, 1)$; the function $y = \tan x$ is bounded in the interval $(0, \frac{\pi}{4})$ and unbounded in the interval $(0, \frac{\pi}{2})$. In connection with this remark we introduce the following definition:

Definition. A function is said to be *bounded* as $x \rightarrow x_0$ if there is a neighbourhood of the point x_0 in which the function is bounded.

Every constant magnitude (interpreted as a function) is of course bounded. Furthermore, every function $y = f(x)$ having a limit as $x \rightarrow x_0$ is bounded as $x \rightarrow x_0$. For, if

$$\lim_{x \rightarrow x_0} f(x) = A$$

then there is a neighbourhood of the point x_0 in which, except for $x = x_0$, the function $f(x)$ is arbitrarily close to the number A .

Every function tending to infinity as x approaches x_0 is of course unbounded as $x \rightarrow x_0$.

It should be noted that the converse is not necessarily true because there are unbounded functions which do not tend to infinity. For instance, the function $y = x \sin x$ assumes arbitrarily large values (we have $y_{x=\frac{\pi}{2}+2k\pi} = \frac{\pi}{2} + 2k\pi$ where k is any integer) and hence it is not bounded as $x \rightarrow \infty$; at the same time it turns into zero for arbitrarily large values of x (since $y_{x=k\pi} = 0$) and therefore cannot tend to infinity. Indeed, if it approached infinity the absolute value of y would be arbitrarily large for all sufficiently large values of x and could not be equal to zero. Let the

reader draw a sketch of the graph of the function $y = x \sin x$ and interpret this argument geometrically.

28. Infinitesimals.

Definition. A function $\alpha(x)$ tending to zero as $x \rightarrow x_0$ is called an *infinitesimal* as $x \rightarrow x_0$.

According to what was indicated, this means that, given any $\varepsilon > 0$ (however small), there is $\delta > 0$ such that the conditions $x \neq x_0$ and $|x - x_0| < \delta$ imply the inequality $|\alpha(x)| < \varepsilon$. Let the reader state the corresponding test for the case $x \rightarrow \infty$. Examples of infinitesimals are the functions

$$\begin{aligned} y &= x^2 \text{ for } x \rightarrow 0, & y &= x-1 \text{ for } x \rightarrow 1 \\ y &= \frac{1}{x} \text{ for } x \rightarrow \infty, & y &= 2^x \text{ for } x \rightarrow -\infty \end{aligned}$$

A concrete nonzero constant number, however small, must not be confused with an infinitesimal. *The only constant number admitting of the interpretation as an infinitesimal is the number zero* (since the limit of a constant is equal to that constant).

Every infinitesimal is of course *bounded* as $x \rightarrow x_0$.

Infinitely large magnitudes and infinitesimals play a very important role in mathematical analysis. The former are functions having no limits while the latter have zero limits; there is a simple relationship between them expressed by the following theorem:

Theorem. If a function $f(x)$ tends to infinity as $x \rightarrow x_0$, then $\frac{1}{f(x)}$ is an infinitesimal; if $\alpha(x)$ is an infinitesimal as $x \rightarrow x_0$, which does not take a zero value for $x \neq x_0$, then $\frac{1}{\alpha(x)}$ is an infinitely large magnitude.

Proof. Let $f(x) \rightarrow \infty$ as $x \rightarrow x_0$. We must show that $\frac{1}{f(x)} \rightarrow 0$ as $x \rightarrow x_0$.

Let us take an arbitrarily small number $\varepsilon > 0$ and the number $M = \frac{1}{\varepsilon}$. Since $f(x)$ is infinitely large, the inequality

$$|f(x)| > M = \frac{1}{\varepsilon}$$

holds for all $x \neq x_0$ lying sufficiently close to x_0 . But then

$$\left| \frac{1}{f(x)} \right| < \frac{1}{M} = \varepsilon$$

which means that $\lim_{x \rightarrow x_0} \frac{1}{f(x)} = 0$. The second part of the theorem is proved analogously under the assumption that the function $\frac{1}{\alpha(x)}$

is defined in a neighbourhood of the point x_0 , that is $\alpha(x) \neq 0$ for the values $x \neq x_0$ belonging to this neighbourhood.

The theorems below are important for applications.

Direct Theorem. If a function has a limit it is representable as a sum of a constant, equal to that limit, and an infinitesimal.

Proof. Let $\lim_{x \rightarrow x_0} f(x) = A$. Then, given an arbitrarily small positive number ε , we have $|f(x) - A| < \varepsilon$ for all $x \neq x_0$ lying sufficiently close to x_0 , which, in accordance with the definition, implies that $f(x) - A$ is an infinitesimal. Consequently,

$$f(x) - A = \alpha(x), \text{ i.e. } f(x) = A + \alpha(x)$$

where $\alpha(x)$ is an infinitesimal as $x \rightarrow x_0$. The function $\alpha(x)$ can of course take on positive, or negative or zero values, that is the function $f(x)$ can be greater than or less than or equal to its limit.

Reciprocal Theorem. If a function is representable as a sum of a constant and an infinitesimal the constant summand is the limit of the function.

Proof. It follows from the equality $f(x) = A + \alpha(x)$ where A is a constant and $\alpha(x)$ an infinitesimal as $x \rightarrow x_0$ that, given an arbitrarily small positive number ε , we have

$$|f(x) - A| = |\alpha(x)| < \varepsilon$$

for all x 's sufficiently close to x_0 . But this exactly means that A is the limit of $f(x)$, i.e.

$$\lim_{x \rightarrow x_0} f(x) = A$$

which is what we set out to prove.

29. General Rules for Finding Limits.

I. Here we shall establish some simple *rules for finding limits* of functions. We shall begin with theorems on infinitesimals and then draw from them, as consequences, the corresponding propositions for functions tending to limits which are not necessarily zero. These theorems provide for the rules facilitating the determination of limits *when given functions are obtained by means of arithmetical operations from functions possessing known limits*.

For brevity, we shall denote functions possessing limits by u , v , w , α , β , etc. and write $\lim u$, $\lim v$ and so on without stipulating in which way the independent variable x changes in the limiting processes under consideration. However, when speaking about several functions we suppose that they tend to their limits in one and the same process of variation of the independent variable (that is their common argument x is supposed to approach one and the same point x_0 or ∞). As before, the proofs will be

given for the case $x \rightarrow x_0$. Let the reader introduce the necessary changes for the case $x \rightarrow \infty$.

Theorem I. A sum of two, three and, generally, of any finite number of infinitesimals is again an infinitesimal.

Proof. Let us take two infinitesimals α and β and prove that their sum

$$\omega = \alpha + \beta$$

is also an infinitesimal.

Choose an arbitrary positive number ε . There exists a neighbourhood of the point x_0 (to which the argument x tends) such that for all its points (except possibly for the point x_0 itself) we have¹

$$|\alpha| < \frac{\varepsilon}{2} \quad \text{and} \quad |\beta| < \frac{\varepsilon}{2}$$

The absolute value of a sum being not greater than the sum of the absolute values, we have

$$|\omega| = |\alpha + \beta| \leq |\alpha| + |\beta|$$

and hence, in that neighbourhood,

$$|\omega| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

which shows that ω is an infinitesimal as $x \rightarrow x_0$. It is apparent that this proof also applies to a *difference* of infinitesimals.

Now let ω be the sum of three infinitesimals α , β and γ :

$$\omega = \alpha + \beta + \gamma$$

This expression can be interpreted as the sum of *two* terms: $\omega = (\alpha + \beta) + \gamma$. According to the above, $(\alpha + \beta)$ is an infinitesimal, and therefore the same argument shows that ω is also an infinitesimal.

Proceeding in this way we readily prove the theorem for any finite number of infinitesimal summands.

The stipulation that the number of summands is *finite* is essentially important. The matter is that in mathematical analysis we also encounter sums of special character in which both the summands and their number vary simultaneously. To sums of this

¹ If the neighbourhoods in which $|\alpha| < \frac{\varepsilon}{2}$ and $|\beta| < \frac{\varepsilon}{2}$ do not coincide both inequalities hold within the smallest of them which can be taken as the neighbourhood we are speaking of.

type the theorem does not apply. If, for example, the number of summands increases indefinitely as each of them tends to zero the assertion of the theorem may not be true. For instance, let

$$\alpha = \frac{1}{n^2}, \quad \beta = \frac{2}{n^2}, \dots, \quad \eta = \frac{n}{n^2}$$

When $n \rightarrow \infty$ each of these quantities tends to zero but their number also increases. The sum

$$\omega = \frac{1+2+\dots+n}{n^2} = \frac{n(n+1)}{2n^2} = \frac{1}{2} + \frac{1}{2n}$$

is not an infinitesimal as $n \rightarrow \infty$ since it tends to $\frac{1}{2}$.

A direct consequence of Theorem I is

Theorem I'. The limit of a sum of a finite number of functions is equal to the sum of the limits of these functions.

Proof. Suppose we are given a definite number k of functions $u, v, \dots, t$ tending, respectively, to their limits $a, b, \dots, d$. We now must take their sum

$$w = u + v + \dots + t$$

and prove that

$$\begin{aligned} \lim w &= \lim (u + v + \dots + t) = \lim u + \lim v + \dots + \lim t = \\ &= a + b + \dots + d \end{aligned}$$

We have (see the direct theorem in Sec. 28)

$$u = a + \alpha, \quad v = b + \beta, \quad \dots, \quad t = d + \tau$$

where $\alpha, \beta, \dots, \tau$ are infinitesimals. Consequently,

$$w = (a + b + \dots + d) + (\alpha + \beta + \dots + \tau)$$

where $\omega = \alpha + \beta + \dots + \tau$ is, by Theorem I, an infinitesimal since it is a sum of a finite number k of infinitesimals. The function w being the sum of the constant $a + b + \dots + d$ and the infinitesimal ω , this constant is the limit of w (see the converse theorem in Sec. 28).

Theorem II. The product of a bounded function by an infinitesimal is an infinitesimal.

Proof. Let α be an infinitesimal and u a bounded function in a neighbourhood of the point x_0 to which x tends: $|u| \leq M$. We must prove that $\lim(u\alpha) = 0$.

Let us take an arbitrary positive number ε . In a neighbourhood of the point x_0 (except possibly at the point x_0 itself) the

values of α satisfy the inequality $|\alpha| < \frac{\varepsilon}{M}$. Consequently¹,

$$|u\alpha| = |u||\alpha| < M \frac{\varepsilon}{M} = \varepsilon$$

whence follows the assertion of the theorem.

In particular, if $c = \text{const}$ and $\alpha \rightarrow 0$ we have

$$\lim (c\alpha) = 0$$

An infinitesimal being always bounded, *the product of two infinitesimals is also an infinitesimal.*

The assertion of the theorem may be incorrect if u is not a bounded function (for instance, $x^2 \cdot \frac{1}{x} \rightarrow \infty$ for $x \rightarrow \infty$).

Theorems I and II imply

Theorem II'. The limit of a product of a finite number of functions is equal to the product of the limits of these functions.

Proof. Retaining the notation introduced in Theorem I' we shall prove that if $w = uv, \dots, t$ then

$$\lim w = \lim (uv, \dots, t) = \lim u \lim v \dots \lim t = ab \dots d$$

We begin with the product of two functions u and v . Since

$$u = a + \alpha \quad \text{and} \quad v = b + \beta$$

we have

$$w = uv = (a + \alpha)(b + \beta) = ab + (a\beta + b\alpha + \alpha\beta)$$

By Theorems I and II, the sum of the last three terms is an infinitesimal, and therefore

$$\lim w = ab$$

Now let there be given three functions u, v and t . Then

$$\lim w = \lim (uv)t = \lim [(uv)t] = \lim (uv) \lim t = \lim u \lim v \lim t$$

Proceeding in this way we easily generalize the theorem to any finite number of factors.

In particular, this theorem implies the following corollaries:

(1) *A constant factor can be taken outside the limit sign:*

$$\lim cu = c \lim u$$

for the limit of a constant is equal to that constant.

¹ Here, as in the footnote on page 86, it is supposed that if there are two different neighbourhoods of the point x_0 where, respectively, $|u| \leq M$ and $|\alpha| < \frac{\varepsilon}{M}$ we take, as the neighbourhood under consideration, the smallest of them.

(2) *The limit of the n th power of a function where n is a positive integer is equal to the n th power of the limit of that function:*

$$\lim u^n = \lim (u \cdot u \dots u) = \lim u \lim u \dots \lim u = (\lim u)^n$$

Theorem III. *The ratio of an infinitesimal to a function having a nonzero limit is an infinitesimal.*

Proof. Let α be an infinitesimal and u a function whose limit is different from zero: $\lim u = A \neq 0$. We have to prove that $\frac{\alpha}{u} \rightarrow 0$.

To this end, let us show that $\frac{1}{u}$ is a bounded function. For definiteness, let $A > 0$. The values of the function u approach the number A as $x \rightarrow x_0$, and there must exist a neighbourhood of x_0 in which they become greater than $\frac{A}{2}$, i.e. $u > \frac{A}{2}$. Then, for the points of this neighbourhood, we have

$$0 < \frac{1}{u} < \frac{2}{A}$$

which means that the function $\frac{1}{u}$ is bounded in that neighbourhood.

In the case $A < 0$ the proof is carried out analogously.

Now, since the quotient $\frac{\alpha}{u}$ can be regarded as the product of the bounded function $\frac{1}{u}$ by the infinitesimal α , Theorem II implies that

$$\lim \frac{\alpha}{u} = 0$$

which is what we had to prove.

Theorem III'. *The limit of a quotient is equal to the quotient of the limits of the dividend and the divisor provided that the limit of the divisor is nonzero.*

Proof. Let $\lim u = a$ and $\lim v = b \neq 0$. Then $u = a + \alpha$ and $v = b + \beta$ where α and β are infinitesimals. We have to show that

$$\lim w = \lim \frac{u}{v} = \frac{\lim u}{\lim v} = \frac{a}{b}$$

Indeed, we have

$$w = \frac{u}{v} = \frac{a + \alpha}{b + \beta} = \frac{a}{b} + \frac{b\alpha - a\beta}{b(b + \beta)}$$

where, according to Theorem III, the fraction $\frac{b\alpha - a\beta}{b(b + \beta)}$ is an infinitesimal, since, by the foregoing theorems, its numerator is an infinitesimal and the limit of the denominator, equal to b^2 , is, by the hypothesis, different from zero. Hence, $\lim w = \frac{a}{b}$.

The theorem becomes inapplicable when $b=0$.

Below is an example demonstrating the given rules:

$$\begin{aligned}\lim_{x \rightarrow 2} \frac{3x^3 - 2x^2 + x + 1}{x^2 - 5x + 3} &= \frac{\lim (3x^3 - 2x^2 + x + 1)}{\lim (x^2 - 5x + 3)} = \\ &= \frac{3(\lim x)^3 - 2(\lim x)^2 + \lim x + 1}{(\lim x)^2 - 5(\lim x) + 3} = \frac{3 \cdot 2^3 - 2 \cdot 2^2 + 2 + 1}{2^2 - 5 \cdot 2 + 3} = -6\frac{1}{3}\end{aligned}$$

In this example we are interested in the limit of a *rational* function as its argument tends to a finite point x_0 at which the denominator of the fraction does not vanish. It turns out that to find such a limit it is sufficient to substitute the number x_0 for the value of the independent variable into the expression of the function¹.

Now we proceed to consider some cases to which the general theorems on limits proved here are inapplicable. This is most often the case when we deal with the limit of a quotient $\frac{u}{v}$ in which the limit of the denominator is equal to zero. If, in these circumstances, the limit of the numerator is different from zero the quotient $\frac{u}{v}$ is an infinitely large magnitude (see Theorem III)

and $\lim \frac{u}{v} = \infty$ while the quotient $\frac{v}{u}$ is an infinitesimal. If both the numerator and the denominator tend to zero simultaneously the determination of the limit requires an additional transformation or a special technique.

Similarly, an additional investigation is required in the cases when the function is not defined at the point x_0 or when the argument x approaches infinity. Let us consider some examples demonstrating these cases.

Examples. (1) Let us find $\lim_{x \rightarrow 3} \frac{x-3}{x^2-9}$. Here, the denominator tends to zero as $x \rightarrow 3$ and the numerator also tends to zero. But since $x^2-9=(x-3)(x+3)$, we have $\frac{x-3}{x^2-9} = \frac{1}{x+3}$ for all the values of x different from $x=3$. Therefore,

$$\lim_{x \rightarrow 3} \frac{x-3}{x^2-9} = \lim_{x \rightarrow 3} \frac{x-3}{(x-3)(x+3)} = \lim_{x \rightarrow 3} \frac{1}{x+3} = \frac{1}{6}$$

In the solution of this problem we cancel by $x-3$, and one may think that this is illegitimate since $x \rightarrow 3$, and the division by zero is not allowed. But this is not the case here: the func-

¹ As will be shown in § 2, this simple technique for determining limits is applicable to every elementary function provided that the point x_0 belongs to the domain of definition of that function.

tions $y = \frac{x-3}{x^2-9}$ and $y = \frac{1}{x+3}$ coincide identically for all $x \neq 3$, and the definition of a limit of a function for $x \rightarrow x_0$ does not involve the value of that function at the point $x = x_0$ itself, and therefore the limits of the above functions as $x \rightarrow 3$ are equal to each other. The essence of this transformation is that the limit of the new function is found easier than that of the original function.

(2) Consider $\lim_{x \rightarrow 1} \frac{2x-3}{x^2-5x+4}$. This function tends to ∞ since the denominator of the fraction is an infinitesimal for $x \rightarrow 1$ while the numerator does not tend to zero. Indeed, we have $\lim_{x \rightarrow 1} (x^2-5x+4) = 0$ and $\lim_{x \rightarrow 1} (2x-3) = -1$; therefore, by Theorem III,

$$\lim_{x \rightarrow 1} \frac{x^2-5x+4}{2x-3} = 0$$

and, since the reciprocal of an infinitesimal is an infinitely large magnitude,

$$\lim_{x \rightarrow 1} \frac{2x-3}{x^2-5x+4} = \infty$$

(3) Now we compute $\lim_{x \rightarrow \frac{1}{2}} \frac{4x^2-4x+1}{4x^2-1}$. Here both the denominator and the numerator are infinitesimals. We have

$$\frac{4x^2-4x+1}{4x^2-1} = \frac{(2x-1)^2}{(2x-1)(2x+1)} = \frac{2x-1}{2x+1}$$

and, consequently,

$$\lim_{x \rightarrow \frac{1}{2}} \frac{4x^2-4x+1}{4x^2-1} = \lim_{x \rightarrow \frac{1}{2}} \frac{2x-1}{2x+1} = \frac{0}{2} = 0$$

(4) Take $\lim_{x \rightarrow \infty} \frac{x+1}{x-1}$. Dividing the numerator and the denominator by x we obtain

$$\lim_{x \rightarrow \infty} \frac{x+1}{x-1} = \lim_{x \rightarrow \infty} \frac{1+\frac{1}{x}}{1-\frac{1}{x}} = \frac{1+\lim_{x \rightarrow \infty} \frac{1}{x}}{1-\lim_{x \rightarrow \infty} \frac{1}{x}} = \frac{1+0}{1-0} = 1$$

(5) If $x \rightarrow \infty$, a rational function tends to zero or to infinity or to a finite nonzero limit depending on whether the degree of the numerator is less than or greater than or equal to the degree of the denominator. Indeed, let us take a rational function

$$y = \frac{a_0 x^m + a_1 x^{m-1} + \dots + a_m}{b_0 x^n + b_1 x^{n-1} + \dots + b_n}$$

and divide the numerator and the denominator by x^n :

$$y = \frac{a_0 x^{m-n} + a_1 x^{m-1-n} + \dots + a_m x^{-n}}{b_0 + b_1 x^{-1} + \dots + b_n x^{-n}}$$

As $x \rightarrow \infty$, the denominator of the transformed fraction tends to b_0 while the numerator tends to zero if $m < n$, to ∞ if $m > n$ and to a_0 if $m = n$. It is clear that the limit is the same for $x \rightarrow +\infty$ and for $x \rightarrow -\infty$. Thus,

$$\lim_{x \rightarrow \infty} y = 0 \quad \text{if } m < n$$

$$\lim_{x \rightarrow \infty} y = \infty \quad \text{if } m > n$$

and

$$\lim_{x \rightarrow \infty} y = \frac{a_0}{b_0} \quad \text{if } m = n$$

II. Next we proceed to consider the rules for passing to the limit in inequalities.

Theorem. If two functions $f(x)$ and $F(x)$ satisfy the inequality

$$f(x) < F(x) \quad (*)$$

for all the values of x belonging to a neighbourhood of a point x_0 and not coinciding with x_0 then

$$\lim_{x \rightarrow x_0} f(x) \leq \lim_{x \rightarrow x_0} F(x) \quad (**)$$

provided that the limits of both functions, as $x \rightarrow x_0$, exist.

Proof. Let $\lim_{x \rightarrow x_0} f(x) = A$ and $\lim_{x \rightarrow x_0} F(x) = B$. We must prove that $A \leq B$. Suppose the contrary, i.e. $A > B$. Then $A - B = C > 0$. Let us take $\varepsilon = \frac{C}{4}$. According to the definition of a limit, there exists a neighbourhood of the point x_0 such that for the values $x \neq x_0$ belonging to that neighbourhood the inequalities

$$-\frac{C}{4} < f(x) - A < \frac{C}{4} \quad \text{and} \quad -\frac{C}{4} < F(x) - B < \frac{C}{4}$$

hold¹. These inequalities imply $f(x) > A - \frac{C}{4}$ and $F(x) < B + \frac{C}{4}$, and hence

$$f(x) - F(x) > A - \frac{C}{4} - B - \frac{C}{4} = \frac{C}{2} > 0$$

which contradicts inequality (*). Thus, the assumption that $A > B$ leads to a contradiction, and therefore $A \leq B$, which is what we had to prove.

¹ Here we take a neighbourhood of the point x_0 such that for the points belonging to it and different from x_0 the inequality (*) and the last two inequalities are fulfilled simultaneously.

Note that a strict inequality connecting functions may lead to a nonstrict inequality for their limits. For example, if $0 < |x| < 1$ then $x^2 < |x|$. But nevertheless,

$$\lim_{x \rightarrow 0} x^2 = \lim_{x \rightarrow 0} |x| = 0$$

The theorem remains of course true if the *strict* inequality $f(x) < F(x)$ entering into its formulation is replaced by the *non-strict* inequality $f(x) \leq F(x)$. In the case $x \rightarrow +\infty$ we must require that inequality (*) should hold for $x > N$ where N is a fixed number.

An important particular case of the theorem arises when one of the two functions $F(x)$ and $f(x)$ is a constant. For definiteness, let $F(x) = \text{const}$. Then the theorem reads:

If $f(x) < M$ (or $f(x) \leq M$) for $0 < |x - x_0| < l$ where l is a fixed number then $\lim_{x \rightarrow x_0} f(x) \leq M$ provided that the limit of the function $f(x)$ exists.

Let the reader state analogous propositions for limits of sequences.

30. Test for the Existence of the Limit of a Function. The Limit of the Function $\frac{\sin x}{x}$ as $x \rightarrow 0$. There are various tests for finding a limit of a function when its direct determination is impractical. Here we confine ourselves to the test given by the following

Theorem. If the values of a function $f(x)$ are contained between the values of two functions $F(x)$ and $\Phi(x)$ tending to one and the same limit A as $x \rightarrow x_0$, the function $f(x)$ also has a limit as $x \rightarrow x_0$ which is equal to the number A .

Proof. Let

$$F(x) \leq f(x) \leq \Phi(x)$$

and

$$\lim_{x \rightarrow x_0} F(x) = \lim_{x \rightarrow x_0} \Phi(x) = A$$

We have to prove that

$$\lim_{x \rightarrow x_0} f(x) = A$$

Take an arbitrary ε -neighbourhood of the number A . By the hypothesis, there exists a δ -neighbourhood of the point x_0 such that for $x \neq x_0$ belonging to this δ -neighbourhood the corresponding values of $F(x)$ and $\Phi(x)$ lie within the ε -neighbourhood of A :

$$A - \varepsilon < F(x) < A + \varepsilon \quad \text{and} \quad A - \varepsilon < \Phi(x) < A + \varepsilon$$

By the inequalities given in the condition of the theorem, the values of $f(x)$ corresponding to the points of that δ -neighbourhood

of the point x_0 are also in the ϵ -neighbourhood of the number A . Since ϵ is arbitrary, this means that

$$\lim_{x \rightarrow x_0} f(x) = A$$

which is what we set out to prove.

Let the reader consider the geometrical meaning of this theorem.

Now we shall apply this test to derive the already mentioned (Sec. 24) limiting relation

$$\lim_{\alpha \rightarrow 0} \frac{\sin \alpha}{\alpha} = 1$$

playing an important role in mathematical analysis.

Theorem. The function $\frac{\sin \alpha}{\alpha}$ possesses the limit equal to 1 as $\alpha \rightarrow 0$:

$$\lim_{\alpha \rightarrow 0} \frac{\sin \alpha}{\alpha} = 1$$

Proof. We shall proceed from the geometrical definition of the sine. Take a circle of unit radius and suppose that the central angle α expressed in radians is contained within the limits 0 and $\frac{\pi}{2}$: $0 < \alpha < \frac{\pi}{2}$.

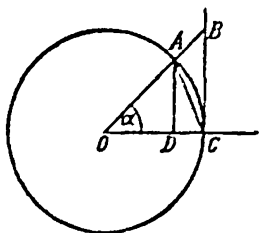


Fig. 42

(Since $\frac{\sin \alpha}{\alpha}$ is an even function, it is sufficient to consider the case when $\alpha > 0$.) Fig. 42 indicates that

$$\text{area of } \triangle OAC < \text{area of sector } OAC < \text{area of } \triangle OBC$$

These areas being respectively equal to $\frac{1}{2} \sin \alpha$, $\frac{1}{2} \alpha$ and $\frac{1}{2} \tan \alpha$, we have

$$\sin \alpha < \alpha < \tan \alpha$$

On dividing all the members of these inequalities by $\sin \alpha$ we obtain

$$1 < \frac{\alpha}{\sin \alpha} < \frac{1}{\cos \alpha}$$

that is,

$$\cos \alpha < \frac{\sin \alpha}{\alpha} < 1$$

But we have $\cos \alpha = OD = 1 - DC > 1 - AC > 1 - \alpha$, and, consequently,

$$1 - \alpha < \frac{\sin \alpha}{\alpha} < 1$$

Now, passing to the limit as $\alpha \rightarrow 0$ we obtain, in accordance with the above test, the desired result:

$$\lim_{\alpha \rightarrow 0} \frac{\sin \alpha}{\alpha} = 1$$

Many other limits can be computed with the aid of this result.
Examples.

$$(1) \lim_{x \rightarrow 0} \frac{\tan x}{x} = \lim_{x \rightarrow 0} \frac{\sin x}{x} \frac{1}{\cos x} = \lim_{x \rightarrow 0} \frac{\sin x}{x} \lim_{x \rightarrow 0} \frac{1}{\cos x} = 1 \cdot 1 = 1$$

$$(2) \lim_{x \rightarrow 0} \frac{\sin kx}{x} = \lim_{x \rightarrow 0} \frac{k \sin kx}{kx} = k \lim_{x \rightarrow 0} \frac{\sin kx}{kx} = k \cdot 1 = k$$

$$(3) \lim_{x \rightarrow 0} \frac{1 - \cos x}{x} = \lim_{x \rightarrow 0} \frac{2 \sin^2 \frac{x}{2}}{x} = \lim_{x \rightarrow 0} \frac{\sin \frac{x}{2}}{\frac{x}{2}} \lim_{x \rightarrow 0} \sin \frac{x}{2} = 1 \cdot 0 = 0$$

31. Test for the Existence of the Limit of a Sequence. The Limit of the Sequence $\left(1 + \frac{1}{n}\right)^n$ as $n \rightarrow \infty$. Consider a function of the integral argument n :

$$y_n = f(n)$$

As was said, it can be interpreted equivalently as the sequence

$$y_1, y_2, \dots, y_n \dots$$

We shall say that this sequence is *bounded above* if there is a number M such that

$$y_n < M$$

for all n . Analogously, we say that the sequence is *bounded below* if there is a number m such that

$$y_n > m$$

for all n . The numbers M and m are called, respectively, an *upper bound* and a *lower bound* of the sequence. A sequence possessing both an upper bound and a lower bound is said to be *bounded*.

We shall additionally suppose that the sequence $y_1, y_2, \dots, y_n, \dots$ is *monotone* (this means that its terms only either increase or decrease as n grows).

Now we state as a theorem an important test for the existence of the limit of a sequence:

Theorem. If $y_1, y_2, \dots, y_n$ is an increasing sequence bounded above it possesses a limit.

The sequence increasing, we have

$$y_1 < y_2 < \dots < y_n < \dots$$

Since it is bounded above there exists an upper bound M : $y_n < M$ for any n . The theorem asserts that under these conditions the sequence has a limit A . Indeed, since the point $y_n = f(n)$ only moves upward along Oy and cannot leave the interval $[0, M]$ it must approach indefinitely a point A^1 ; according to Sec. 29, II, $A \leq M$.

If a sequence $y_n = f(n)$ is monotone and unbounded then $\lim_{n \rightarrow \infty} f(n) = +\infty$, that is, the point y_n , when moving upward along Oy , leaves any neighbourhood of the point $y = 0$.

The situation is similar in the case of a decreasing sequence

$$y_1 > y_2 > \dots > y_n > \dots$$

If such a sequence is bounded below, that is $y_n > m$ for any n , it has a limit A which is not less than m : $A \geq m$. If the sequence is unbounded, the function $f(n)$ approaches $-\infty$ as $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} f(n) = -\infty$$

It should be noted that for a nonmonotone sequence there are not *two* but *three* possible cases:

- (1) the sequence has a limit;
- (2) the sequence tends to infinity;
- (3) the sequence has neither finite nor infinite limit (for example, such is the sequence $y_n = (-1)^n$).

The above theorem sometimes makes it possible to establish the existence of a limit although it does not specify its numerical value. We shall apply it to prove the existence of a limit which plays an extremely important role in mathematical analysis. This result is expressed by the following theorem:

Theorem. The sequence $y_n = \left(1 + \frac{1}{n}\right)^n$ possesses a limit as $n \rightarrow \infty$.

Proof. By Newton's binomial formula, we have²

$$\begin{aligned} y_n &= \left(1 + \frac{1}{n}\right)^n = 1 + \frac{n}{1!} \frac{1}{n} + \frac{n(n-1)}{2!} \frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!} \frac{1}{n^3} + \dots \\ &\dots + \frac{n(n-1)\dots(n-n+1)}{n!} \frac{1}{n^n} = 1 + 1 + \frac{1}{2!} \left(1 - \frac{1}{n}\right) + \\ &+ \frac{1}{3!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) + \dots + \frac{1}{n!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{n-1}{n}\right) \end{aligned}$$

¹ Here we do not present the rigorous proof of this proposition confining ourselves to the intuitive argument.

² The last binomial coefficient, that is the factor in $\frac{1}{n^n}$, is of course equal to 1; for the sake of convenience it is represented here in accordance with the general formula for the binomial coefficients. For the binomial formula itself we refer the reader to textbooks on algebra.

Replacing n by $n+1$ we obtain an analogous expression for y_{n+1} . Next, comparing these expressions we conclude that the first two terms in both sums coincide and that every subsequent term in y_n is less than the corresponding term in y_{n+1} (since every expression in the parentheses $(1-\frac{1}{n})$, $(1-\frac{2}{n})$, etc. becomes greater as n increases); finally, the expression of y_{n+1} contains an additional (the last) positive term. Therefore it is apparent that $y_n < y_{n+1}$ for any n , that is, y_n is an increasing sequence. Let us show that it is bounded. To this end, we replace the proper fractions in the parentheses by unities and thus receive

$$y_n < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}$$

The right-hand side of this inequality still increases if we replace

$$\frac{1}{3!} = \frac{1}{2 \cdot 3} \quad \text{by} \quad \frac{1}{2 \cdot 2} = \frac{1}{2^2}$$

$$\frac{1}{4!} = \frac{1}{2 \cdot 3 \cdot 4} \quad \text{by} \quad \frac{1}{2 \cdot 2 \cdot 2} = \frac{1}{2^3}$$

$$\frac{1}{n!} = \frac{1}{2 \cdot 3 \cdots n} \quad \text{by} \quad \frac{1}{2 \cdot 2 \cdots 2} = \frac{1}{2^{n-1}}$$

Consequently,

$$y_n < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^{n-1}}$$

Now, adding to the right-hand expression the terms $\frac{1}{2^n}, \frac{1}{2^{n+1}}, \dots$, we replace it by unity plus the infinite geometric progression with common ratio $\frac{1}{2}$:

$$y_n < 1 + \left(1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^n} + \dots\right)$$

The sum of the progression in the parentheses being equal to 2, we obtain

$$y_n < 3$$

Thus, the increasing sequence in question is bounded above, and hence, according to the above test, it possesses a finite limit. This limit plays a very important role in mathematics and is traditionally denoted by e .¹

¹ This notation was introduced by L. Euler (1707-1783) who, by the way, also introduced the notation π for the ratio of the circumference of a circle to its diameter. L. Euler (a Swiss by birth) was a great scientist. He spent most of his life in Russia and died in St. Petersburg. L. Euler contributed many outstanding results to mathematical analysis, celestial mechanics, shipbuilding and other divisions of science.

Definition. The number e is the limit

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

The number e is irrational and therefore cannot be expressed precisely by a finite decimal. Its approximate value is

$$e \approx 2.718$$

It can be shown (but we do not present the proof here) that the function

$$y = \left(1 + \frac{1}{x}\right)^x$$

has a limit not only when its argument runs through natural numbers (that is $x=n$ where $n=1, 2, 3, \dots$) but also when it varies continuously and approaches $-\infty$ or $+\infty$. In all the cases the limit is the same number e .

Thus,

$$\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e$$

Here we do not present the general proof of this assertion.

The test given in this section can be generalized without any changes to functions of a continuous argument: *if a function $f(x)$ increases as $x \rightarrow +\infty$ and remains bounded (see Sec. 27, II) it possesses a limit.* We suggest that the reader construct examples showing that if at least one of these conditions (that is the monotonicity or the boundedness) is violated the function may have no limit. Analogous tests can be stated for a decreasing function and for the case $x \rightarrow -\infty$.

§ 2. Continuous Functions

32. Continuity. Recall that the *increment* of a function $y=f(x)$ at a given point x_0 is the difference

$$\Delta y = f(x_0 + \Delta x) - f(x_0)$$

where Δx is the increment of the argument (Fig. 43). Now we state the following definition:

Definition. A function $y=f(x)$ is said to be *continuous at a point x_0* if it is defined in a neighbourhood of the point x_0 and

$$\lim_{\Delta x \rightarrow 0} \Delta y = 0$$

that is, if to an infinitesimal increment of the argument there corresponds an infinitesimal increment of the function¹.

For instance, the function $y = x^3$ is continuous at every point x_0 . For, since

$$\Delta y = (x_0 + \Delta x)^3 - x_0^3 = 3x_0^2\Delta x + 3x_0\Delta x^2 + \Delta x^3$$

it is clear that if $\Delta x \rightarrow 0$ then also $\Delta y \rightarrow 0$, which means that the function is continuous. (Let the reader prove that the functions $\sin x$ and $\cos x$ are continuous at every point x_0 .)

Using the expression of Δy we can rewrite the condition of continuity in the form

$$\lim_{\Delta x \rightarrow 0} [f(x_0 + \Delta x) - f(x_0)] = 0$$

or, equivalently,

$$\lim_{\Delta x \rightarrow 0} f(x_0 + \Delta x) = f(x_0)$$

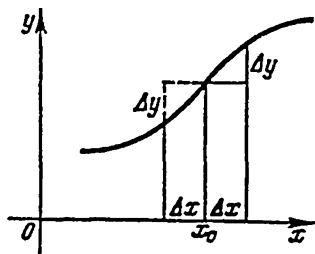


Fig. 43

If $x_0 + \Delta x$ is denoted by x the variable x tends to x_0 as $\Delta x \rightarrow 0$ and vice versa. Therefore the latter condition can be rewritten as

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

Hence, we can say:

A function $y = f(x)$ is continuous at a point x_0 if it is defined in a neighbourhood of that point² and if the limit of the function, as the independent variable x tends to x_0 , exists and is equal to the particular value of the function at the point $x = x_0$:

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) \quad (*)$$

This means that, given any $\epsilon > 0$, there is $\delta > 0$ such that for all the values of x satisfying the condition $|x - x_0| < \delta$ the inequality $|f(x) - f(x_0)| < \epsilon$ is fulfilled. The condition $x \neq x_0$ is not stipulated here since the latter inequality also holds for $x = x_0$.

It should be noted that if the value of a function $f(x)$ at a point x_0 where it is continuous is different from zero there exists a neighbourhood of the point x_0 at whose all points the sign of the function $f(x)$ coincides with that of $f(x_0)$. For, because of the

¹ In this definition we speak about the increment $\Delta y = f(x_0 + \Delta x) - f(x_0)$ of the function, and therefore the fact that the function is defined not only in a neighbourhood of the point x_0 but also at the point x_0 itself is implied automatically although this condition is not stipulated explicitly in the statement of the definition.—Tr.

² Of course, here again the function must be defined at the point $x = x_0$ itself.—Tr.

continuity, there is a neighbourhood of the point x_0 in which the difference between the function and its limit (which is equal to $f(x_0)$) is so small that the condition $f(x_0) > 0$ ($f(x_0) < 0$) implies the positivity (negativity) of $f(x)$.

Definition. If a function is continuous at every point of an interval it is said to be continuous in that interval.

In the application to a closed interval the definition of continuity should be slightly changed for the end points of the interval. Namely, for the *left* end point the argument x should only be given *positive* increments Δx and for the *right* end point only *negative* ones.

The condition of continuity of a function in an interval can be described as the property of the function to change "gradually" within that interval in the sense that *small* variations of the argument generate *small* variations of the function itself. This is a characteristic feature of many phenomena and processes. For instance, when a rod is heated we consider it as elongating continuously, the growth of an organism is regarded as a continuous process, the air temperature varies continuously during the day, etc.

The continuity of a function in an interval can be interpreted geometrically as the continuity of its graph which must not have "interruptions".

By the way, when describing the properties and the graphs of the basic elementary functions in Chapter I, we meant tacitly that they were continuous. It can be shown (for the proof we refer the reader to advanced courses in mathematical analysis) that all the basic elementary functions are continuous in the intervals where they are defined. In Sec. 34 we shall extend this general proposition to all elementary functions.

33. Points of Discontinuity of a Function. The examination of the graph of the function $y = \frac{1}{x}$ in the vicinity of the point $x = 0$ clearly shows that it "splits" into two separate curves at that point. The same applies to the graph of the function $y = \tan x$ in the vicinity of the points $x = (2k+1)\frac{\pi}{2}$ (see Sec. 21, Fig. 27) and to the graph of the function $y = \frac{|x|}{x}$ (or $y = \operatorname{sgn} x$) near the point $x = 0$ (see Sec. 24, Fig. 37). It is said that at each point of that kind the function is *discontinuous*.

Definition. If, for a function $y = f(x)$, the condition of continuity, at a point x_0 ,

$$\lim_{x \rightarrow x_0} f(x) = f(x_0) \quad (*)$$

is violated, x_0 is called a *point of discontinuity* of the function.

The condition of continuity is violated if one of the following two cases takes place:

1. The limit entering into condition (*) is infinite or does not exist at all (a function must not necessarily be defined at its point of discontinuity).

2. The limit $\lim_{x \rightarrow x_0} f(x)$ exists but does not coincide with $f(x_0)$.

It should also be noted that if x_0 is an end point of the interval where the function in question is defined the limit $\lim_{x \rightarrow x_0} f(x)$ is understood in the sense that x is only given the values belonging to the interval.

In many problems we encounter the so-called *infinite discontinuities* which appear at points of discontinuity x_0 for which $\lim_{x \rightarrow x_0} f(x) = \infty$. Examples of such points are the point $x=0$ for the

function $y = \frac{1}{x}$, the points $x = (2k+1)\frac{\pi}{2}$ for the function $y = \tan x$, the point $x=0$ for the function $y = \log x$ (which is only defined on the right of that point of discontinuity $x=0$) and the points $x=-1$ and $x=1$ for the function $y = \frac{1}{\sqrt{1-x^2}}$ (in the latter case

the points of discontinuity are the end points of the interval in which the function is defined).

An important class of points of discontinuity is formed by the so-called points of discontinuity of the *first kind*. To state their definition we introduce the notions of a *left-hand* and a *right-hand* limits of a function.

Let x approach a point x_0 so that it always remains on the *left* of x_0 , that is $x < x_0$. If, under this condition, the function $f(x)$ tends to a limit we speak of the *left-hand* limit of the function $f(x)$ at the point x_0 . A *right-hand* limit is defined similarly. The symbols for left-hand and right-hand limits are, respectively, $f(x_0-0)$ and $f(x_0+0)$. Thus,

$$\lim_{\substack{x \rightarrow x_0 \\ x < x_0}} f(x) = f(x_0-0) \quad \text{and} \quad \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} f(x) = f(x_0+0)$$

It appears evident that if a function possesses a limit as its argument x approaches x_0 in an *arbitrary* way the left-hand and the right-hand limits also exist and coincide with the limit of the function. Conversely, if a function has coinciding left-hand and right-hand limits at a point x_0 it tends to the same limit as the argument approaches x_0 arbitrarily.

When we spoke about the condition of continuity (or discontinuity) for an end point of the interval of definition of a function the corresponding limit was in fact understood as the right-hand

limit for the left end point and as the left-hand limit for the right end point.

Let us state the following definition:

Definition. If a function $f(x)$ possesses left-hand and right-hand limits at its point of discontinuity x_0 this point is called a *point of discontinuity of the first kind*.

The most important type of a point of discontinuity of the first kind is the one for which the left-hand and the right-hand limits at that point of discontinuity are distinct. This situation is illustrated geometrically in Fig. 44. As an example, we can take the point $x=0$ for the function $y = \frac{|x|}{x}$ (see the

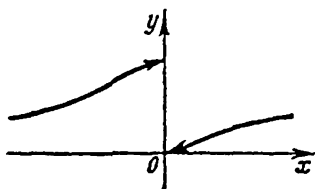


Fig. 44

example at the end of Sec. 24). Another example is the function $y = \arctan \frac{1}{x}$. It is not defined at the point $x=0$. If $x \rightarrow 0$ and remains *negative* the expression $\frac{1}{x}$ tends to $-\infty$ and $y \rightarrow -\frac{\pi}{2}$ (see Sec. 22). If $x \rightarrow 0$ and remains *positive*

we have $\frac{1}{x} \rightarrow +\infty$ and $y \rightarrow \frac{\pi}{2}$.

Thus, $x=0$ is a point of discontinuity of the first kind for the function $y = \arctan \frac{1}{x}$. The sketch of the graph of this function is shown in Fig. 45.

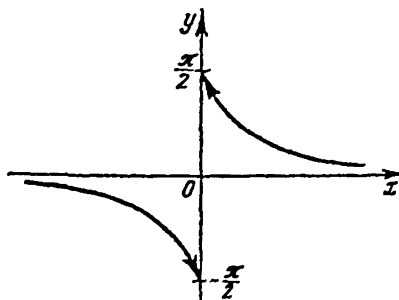


Fig. 45

Now suppose that at a point x_0 a function possesses both left-hand and right-hand limits which coincide: $f(x_0-0) = f(x_0+0)$. If the function $f(x)$ is defined at the point x_0 and $f(x_0)$ is equal to these limits the function is continuous at x_0 . But if $f(x)$ is not defined for $x=x_0$ we say that the function $f(x)$ has a *removable singularity* at the point $x=x_0$. The meaning of this term is explained as follows: if the point x_0 is added to the domain of definition of the function $f(x)$ and if at that new point the value of the function is put equal to the common value of the left-hand and right-hand limits the (new) function $f(x)$ thus obtained is continuous at the point x_0 .

For example, the function $y = \frac{\sin x}{x}$ is not defined at the point $x=0$. But since $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$ we can introduce a new function,

defined for all the values of x and coinciding with the old one for $x \neq 0$, which is everywhere continuous:

$$y = \begin{cases} \frac{\sin x}{x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0 \end{cases}$$

The term "removable singularity" is also applied when a function $f(x)$ is defined at the point x_0 and possesses coinciding left-hand and right-hand limits which are not equal to $f(x_0)$. For instance, such is the point $x=0$ for the function

$$y = f(x) = \begin{cases} x & \text{for } x \neq 0 \\ 2 & \text{for } x = 0 \end{cases}$$

Its graph is obtained from the graph of the continuous function $y=x$ if its point $(0, 0)$ is "torn out" and moved 2 units of length upward along Oy .

All the points of discontinuity not belonging to the class of the points of discontinuity of the first kind are referred to as the *points of discontinuity of the second kind*.

One must not think that a point of discontinuity of the second kind is necessarily a point of infinite discontinuity. There are *bounded* functions having neither a left-hand nor a right-hand limit as the argument approaches a point of discontinuity. As an example of that kind, consider the function $y = \sin \frac{1}{x}$: when x tends to zero the function does not tend to any limit, finite or infinite, left-hand or right-hand. To verify this, construct a sketch of its graph by plotting the points with the coordinates $x = \frac{1}{k\pi}$ and $\frac{1}{(k + \frac{1}{2})\pi}$ where k are integers.

34. Operations on Continuous Functions. Continuity of Elementary Functions. We proceed to prove that, under certain conditions, a finite number of arithmetical operations and operations of constructing a function of a function performed on continuous functions result into a new continuous function. All the propositions below are proved in like manner: we show in each case that the limit of the resultant function is equal to its particular value at the corresponding point, which implies the continuity. For brevity, instead of $u(x)$, $v(x)$, ... we shall write u , v , ...

Theorem I. The sum of a finite number of functions continuous at a point is a continuous function at that point.

Proof. Suppose we are given a definite number of functions u , v , ..., t continuous at a point x_0 . We have to prove that their sum $w = u + v + \dots + t$ is a continuous function at that point. The

functions $u, v, \dots, t$ being continuous, we have

$$\lim_{x \rightarrow x_0} u = u_0, \quad \lim_{x \rightarrow x_0} v = v_0, \quad \dots, \quad \lim_{x \rightarrow x_0} t = t_0$$

where $u_0, v_0, \dots, t_0$ are, respectively, the values of the functions $u, v, \dots, t$ at the point x_0 . By the theorem on the limit of a sum (Sec. 29), we write

$$\begin{aligned} \lim_{x \rightarrow x_0} w &= \lim_{x \rightarrow x_0} (u + v + \dots + t) = \lim_{x \rightarrow x_0} u + \lim_{x \rightarrow x_0} v + \dots + \lim_{x \rightarrow x_0} t = \\ &= u_0 + v_0 + \dots + t_0 = w_0 \end{aligned}$$

where w_0 is the value of the function w at the point $x = x_0$. Thus, $\lim_{x \rightarrow x_0} w = w_0$, which is what we wished to prove.

Theorem II. The product of a finite number of functions continuous at a point is a continuous function at that point.

Proof. Let $w = u \cdot v \dots t$. Retaining the notation of Theorem I and using the theorem on the limit of a product (Sec. 29), we receive

$$\lim_{x \rightarrow x_0} w = \lim_{x \rightarrow x_0} (u \cdot v \dots t) = \lim_{x \rightarrow x_0} u \cdot \lim_{x \rightarrow x_0} v \dots \lim_{x \rightarrow x_0} t = u_0 v_0 \dots t_0 = w_0$$

which is what we set out to prove.

Theorem III. The quotient of two functions continuous at a point x_0 is a continuous function at the point x_0 provided that the denominator does not turn into zero at that point.

Proof. If $w = \frac{u}{v}$, then, by the theorem on the limit of a quotient (Sec. 29), we have

$$\lim_{x \rightarrow x_0} w = \lim_{x \rightarrow x_0} \frac{u}{v} = \frac{\lim_{x \rightarrow x_0} u}{\lim_{x \rightarrow x_0} v} = \frac{u_0}{v_0} = w_0$$

if $\lim_{x \rightarrow x_0} v = v_0 \neq 0$. Therefore $w = \frac{u}{v}$ is a continuous function at the point x_0 .

Theorem IV. A function of a function composed of a finite number of continuous functions is a continuous function.

Proof. It is sufficient to prove this assertion for a function of a function formed of two continuous functions because after that it can be extended, consecutively, to an arbitrary number of constituent functions. Let

$$y = f(z), \quad z = \varphi(x) \quad \text{and} \quad y = f[\varphi(x)] = F(x)$$

where $\varphi(x)$ is continuous for $x = x_0$ and $f(z)$ is continuous for $z = z_0 = \varphi(x_0)$. We have to prove that y as function of x (that is the function $y = F(x)$) is continuous at the point x_0 .

Indeed, let $x \rightarrow x_0$. The continuity of the function $z = \varphi(x)$ implies that $\lim_{x \rightarrow x_0} \varphi(x) = \varphi(x_0) = z_0$, that is $z \rightarrow z_0$. The function $f(z)$ being continuous at the point z_0 , we have $\lim_{z \rightarrow z_0} f(z) = f(z_0)$. Now, since $z = \varphi(x)$, we can rewrite the last relation in the form

$$\lim_{x \rightarrow x_0} f[\varphi(x)] = f[\varphi(x_0)]$$

or, equivalently,

$$\lim_{x \rightarrow x_0} F(x) = F(x_0)$$

which is what we wished to prove.

We also state, without proof, the following theorem:

Theorem V. The inverse of a monotone¹ continuous function is continuous in the interval where it is defined.

As was mentioned, all the basic elementary functions are continuous in the intervals where they are defined (Sec. 32), and therefore the four theorems proved here imply that *every elementary function is continuous in those intervals where it is defined.*

An elementary function can only be discontinuous at those points where some of the constituent functions is formed of are not defined or where the denominators of some fractions involved vanish. For instance, the function $y = \frac{x}{x^2 - 4}$ is discontinuous at the points $x = \pm 2$ and continuous at all the other points; the function $y = x^2 \tan x$ is discontinuous at the points $x = (2k+1)\frac{\pi}{2}$ ($k = 0, \pm 1, \pm 2, \dots$).

It should be noted that the limiting relation $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ expressing the continuity of the function $f(x)$ at the point $x = x_0$ can be rewritten as

$$\lim_{x \rightarrow x_0} f(x) = f(\lim_{x \rightarrow x_0} x) \quad (*)$$

which means that *the symbol of the limit and the symbol of a continuous function can be interchanged.*

It follows that the passage to the limit in an elementary function can be carried out according to the following simple and convenient practical rule:

To find the limit of an elementary function as its argument tends to a value belonging to the domain of definition of the function we should substitute this value for the argument into the expression of the function.

This rule is very important since in applications of mathematical analysis we usually deal with elementary functions.

¹ The monotonicity is understood here not in the broad but in the strict sense (see pp. 40, 41).

The cases when the argument tends to infinity or to a point not belonging to the domain of definition of the function always require a special investigation. The following general rule for passing to the limit (which we give without proof) is sometimes of use in such cases:

If the limit of a function $\varphi(x)$ as $x \rightarrow x_0$ or as $x \rightarrow \infty$ exists and is equal to b and the function $f(x)$ is continuous at the point $x=b$ then

$$\lim_{\substack{x \rightarrow x_0 \\ \text{or} \\ x \rightarrow \infty}} f[\varphi(x)] = f[\lim_{\substack{x \rightarrow x_0 \\ \text{or} \\ x \rightarrow \infty}} \varphi(x)] \quad (**)$$

Examples. (1) Let us find $\lim_{x \rightarrow 0} y$ where $y = \frac{x^3 + 3x^2 - x - 3}{x^2 + x - 6}$. Since the given function is continuous at the point $x=0$ (because the polynomial in the denominator does not vanish at that point) we obtain

$$\lim_{x \rightarrow 0} \frac{x^3 + 3x^2 - x - 3}{x^2 + x - 6} = \frac{0^3 + 3 \cdot 0^2 - 0 - 3}{0^2 + 0 - 6} = \frac{-3}{-6} = \frac{1}{2}$$

Now let us find $\lim_{x \rightarrow -3} y$. At the point $x=-3$ both the numerator and the denominator turn into zero. We have $x^2 + x - 6 = (x+3)(x-2)$ and $x^3 + 3x^2 - x - 3 = (x+3)(x^2 - 1)$, and hence, on cancelling by $(x+3)$, we obtain

$$\lim_{x \rightarrow -3} \frac{x^3 + 3x^2 - x - 3}{x^2 + x - 6} = \lim_{x \rightarrow -3} \frac{x^2 - 1}{x - 2} = \frac{(-3)^2 - 1}{-3 - 2} = -\frac{8}{5}$$

In the same manner we find

$$\lim_{x \rightarrow 2} \frac{x^3 + 3x^2 - x - 3}{x^2 + x - 6} = \infty, \quad \lim_{x \rightarrow 1} \frac{x^3 + 3x^2 - x - 3}{x^2 + x - 6} = 0$$

$$\lim_{x \rightarrow \infty} \frac{x^3 + 3x^2 - x - 3}{x^2 + x - 6} = \infty$$

(2) Let $y = \frac{x^2 - 4}{\cos \frac{\pi}{4} x}$. We shall compute $\lim_{x \rightarrow 1} y$. The given func-

tion is continuous at the point $x=1$, and hence $\lim_{x \rightarrow 1} \frac{x^2 - 4}{\cos \frac{\pi}{4} x} = \frac{1^2 - 4}{\cos \frac{\pi}{4} \cdot 1} = -3\sqrt{2}$. But the limit of this function as $x \rightarrow 2$ can-

not be found by means of the general rule since the denominator turns into zero for $x=2$. Noting that the numerator also turns into zero at the point $x=2$, we put $x-2=z$, i.e. $x=z+2$, and

reduce the case $x \rightarrow 2$ to $z \rightarrow 0$:

$$\lim_{x \rightarrow 2} \frac{x^2 - 4}{\cos \frac{\pi}{4} x} = \lim_{z \rightarrow 0} \frac{z(z+4)}{\cos \frac{\pi}{4} (z+2)} = \lim_{z \rightarrow 0} \frac{z(z+4)}{-\sin \frac{\pi}{4} z}$$

Continuing the computations we find the desired limit:

$$\lim_{z \rightarrow 0} \frac{z(z+4)}{-\sin \frac{\pi}{4} z} = -\frac{4}{\pi} \lim_{z \rightarrow 0} \frac{\frac{\pi}{4} z}{\sin \frac{\pi}{4} z} \lim_{z \rightarrow 0} (z+4) = -\frac{16}{\pi}$$

$$(3) \lim_{x \rightarrow \infty} \ln \left(1 + \frac{1}{x} \right)^x = \ln \left[\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x} \right)^x \right] = \ln e = 1. \text{ Here the}$$

rule (*) is inapplicable, and therefore we use (**).

35. **Properties of Continuous Functions.** A function continuous in a *closed* interval possesses a number of important properties. Here we present, without proof, two of them.

Theorem I. If a function is continuous in a *closed* interval there exist at least one point at which the function assumes the greatest value and at least one point at which it assumes the least value on that interval.

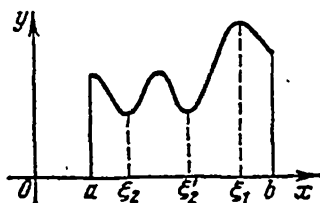


Fig. 46

Let $y = f(x)$ be a continuous function in an interval $[a, b]$ (Fig. 46).

The theorem asserts that there are at least one point $\xi_1 \in [a, b]$ and at least one point $\xi_2 \in [a, b]$ such that at these points the function attains, respectively, its greatest and least values¹:

$$f(x) \leq f(\xi_1) \text{ and } f(x) \geq f(\xi_2) \text{ for all } x \in [a, b]$$

There may be of course several such points; for example, in Fig. 46 we see that the function $f(x)$ takes on its least value at two points (denoted by ξ_1 and ξ_2).

If the condition that the interval is *closed* is dropped the assertion of the theorem may become incorrect. As an example, consider the function $y = x$ which is continuous in every *open* interval (a, b) ; it does not attain the least and the greatest values on such an interval (if the function $y = x$ is regarded as defined in the *closed* interval $[a, b]$ it attains its least and greatest values at the points a and b which are not included into the open interval (a, b)). Similarly, the theorem is inapplicable to discontinuous functions. Let the reader construct an example of a discontinuous function (or draw its graph) which is defined in a

¹ The inclusion relation of the type $\xi \in [a, b]$ means that $a \leq \xi \leq b$ (see Sec. 4).

closed interval and does not achieve its least and greatest values.

Theorem II. If a function defined in a *closed* interval assumes values of different signs at its end points there exists at least one point lying inside the interval at which the function turns into zero.

Suppose that a function $y=f(x)$ is continuous in an interval $[a, b]$, and that $f(a) < 0$ and $f(b) > 0$ (Fig. 47). The theorem asserts that there is at least one point $\xi \in (a, b)$ such that $f(\xi) = 0$. This fact is quite obvious from the geometrical point of view:

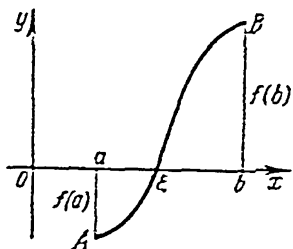


Fig. 47

the graph of the function which is a continuous line joins the point A lying below Ox to the point B above Ox , and therefore it must at least once intersect the axis Ox .

Theorem II can be restated in a more general form:

A function continuous in a closed interval assumes at least once any value lying between its end point values.

Let $f(a) = m$, $f(b) = M$ and $m < M$ (the case $m > M$ is treated quite similarly). Take an arbitrary number μ lying between m and M : $m < \mu < M$. The theorem asserts that there is at least one point $\xi \in (a, b)$ such that $f(\xi) = \mu$. This assertion is readily reduced to Theorem II. Indeed, let us consider the auxiliary function

$$\varphi(x) = f(x) - \mu$$

It assumes the values of different signs $\varphi(a) = f(a) - \mu = m - \mu < 0$ and $\varphi(b) = f(b) - \mu = M - \mu > 0$ at the end points.

The function $\varphi(x)$ being continuous in the interval $[a, b]$, Theorem II implies that there is a point $\xi \in (a, b)$ at which the function vanishes. Thus,

$$\varphi(\xi) = f(\xi) - \mu = 0$$

whence $f(\xi) = \mu$.

Geometrically, this proposition means that any straight line $y = \mu$ parallel to Ox and passing between the initial and the terminal points of a continuous line intersects the line at least once. The abscissa of the point of intersection is the value $x = \xi$ for which $f(\xi) = \mu$ (see Fig. 48 where there are three such points). It follows that

if a continuous function passes from one of its values to another it necessarily passes through all the intermediate values.

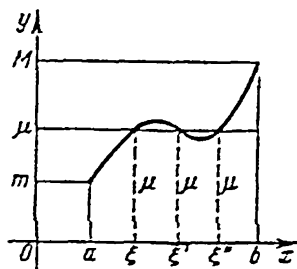


Fig. 48

In particular, a function continuous in an interval assumes at least once every value lying between its least and greatest values.

It is also clear, geometrically, that if a continuous function is increasing (decreasing) in the given closed interval it assumes the least value at the left (right) end point of the interval and the greatest value at the right (left) end point; besides, every intermediate value is assumed exactly once.

§ 3. Comparison of Infinitesimals. Comparison of Infinitely Large Magnitudes

36. Comparison of Infinitesimals. Equivalent Infinitesimals.

I. Comparison of infinitesimals. Let $\alpha(x)$ and $\beta(x)$ be infinitesimals as $x \rightarrow x_0$. To compare the infinitesimals $\alpha(x)$ and $\beta(x)$ means to determine the limit of their ratio

$$\lim_{x \rightarrow x_0} \frac{\alpha(x)}{\beta(x)}$$

provided that it exists. In this section we always assume that the function appearing in the denominator is different from zero for every point $x \neq x_0$ in a neighbourhood of the point x_0 . All the definitions below are readily extended to the case when x tends to ∞ instead of a finite value x_0 .

Let us consider different cases occurring when infinitesimals are compared.

(a) Let $\lim_{x \rightarrow x_0} \frac{\alpha(x)}{\beta(x)} = 0$. Then $\alpha(x)$ is called an *infinitesimal of higher order than* $\beta(x)$.

In this case we write

$$\alpha(x) = o(\beta(x)) \text{ for } x \rightarrow x_0$$

(read " $\alpha(x)$ is small o of $\beta(x)$ for $x \rightarrow x_0$ ").

It should be noted that this symbolic relation is not an equality in the ordinary sense; it only indicates that the limit of the ratio of $\alpha(x)$ to $\beta(x)$ is equal to zero.

Here are examples: $x^2 = o(x)$ for $x \rightarrow 0$ since $\lim_{x \rightarrow 0} \frac{x^2}{x} = \lim_{x \rightarrow 0} x = 0$; similarly $x^3 = o(x)$ for $x \rightarrow 0$ since $\lim_{x \rightarrow 0} \frac{x^3}{x} = 0$.

If $\alpha(x) = o(\beta(x))$ the ratio $\frac{\alpha(x)}{\beta(x)} = \varepsilon(x)$ is an infinitesimal as $x \rightarrow x_0$ and

$$\alpha(x) = \varepsilon(x) \beta(x)$$

The converse is also obvious: if $\alpha(x) = \varepsilon(x)\beta(x)$ where $\varepsilon(x)$ is an infinitesimal and $\beta(x)$ is not equal to zero for $x \neq x_0$ the ratio $\frac{\alpha(x)}{\beta(x)}$ is an infinitesimal, that is $\alpha(x) = o(\beta(x))$.¹

For brevity, we shall not write the argument of the functions $\alpha(x)$, $\beta(x)$, $\gamma(x)$, ... and shall not indicate that $x \rightarrow x_0$ (or $x \rightarrow \infty$). But of course it is always supposed that α , β , γ , ... are functions of one and the same argument and are infinitesimals as the independent variable tends to a point x_0 (or to ∞).

The following rules for the symbol "o" are easily checked:

If $\alpha = o(\gamma)$ then $C\alpha = o(\gamma)$ where C is any constant.

If $\alpha = o(\gamma)$ and $\beta = o(\gamma)$ then $\alpha \pm \beta = o(\gamma)$.

If $\alpha = o(\beta)$ and $\beta = o(\gamma)$ then $\alpha = o(\gamma)$.

If x is an infinitesimal then in the sequence of magnitudes x , x^2 , x^3 , ... every term with a greater exponent is an infinitesimal of higher order than any term with a smaller exponent. This accounts for the origination of the term "of higher order".

The function identically equal to zero is apparently an infinitesimal of higher order than any other infinitesimal, and therefore it is not compared with other functions.

(b) Let $\lim_{x \rightarrow x_0} \frac{\alpha(x)}{\beta(x)} = 1$. Then the infinitesimals $\alpha(x)$ and $\beta(x)$ are said to be *equivalent*.

For equivalent infinitesimals α and β write $\alpha \sim \beta$. It is clear that if $\alpha \sim \beta$ then $\beta \sim \alpha$, that is $\lim \frac{\alpha}{\beta} = 1$ implies $\lim \frac{\beta}{\alpha} = 1$.

Important examples of equivalent infinitesimals are $\sin x$ and x and also $\tan x$ and x for $x \rightarrow 0$ (see Sec. 30).

The following relations are readily verified:

If $\alpha \sim \beta$ and $\beta \sim \gamma$ then $\alpha \sim \gamma$.

If $\alpha = o(\gamma)$ and $\beta \sim \alpha$ then $\beta = o(\gamma)$.

If $\alpha = o(\gamma)$ and $\gamma \sim \delta$ then $\alpha = o(\delta)$.

As an instance, let us prove the second relation. Since $\frac{\beta}{\gamma} = \frac{\beta}{\alpha} \cdot \frac{\alpha}{\gamma}$, and the limit of the first factor on the right is equal to 1 while the limit of the second factor is equal to 0 we have $\lim \frac{\beta}{\gamma} = 0$.

The following theorem describes an important property of equivalent infinitesimals.

¹ The relation $\alpha(x) = \varepsilon(x)\beta(x)$ can of course be taken as an original definition of the symbolic relation $\alpha(x) = o(\beta(x))$.

Theorem. For two infinitesimals $\alpha(x)$ and $\beta(x)$ to be equivalent it is necessary and sufficient that their difference be an infinitesimal of higher order than $\alpha(x)$ (or than $\beta(x)$). (It is supposed that $\alpha(x)$ and $\beta(x)$ do not turn into zero for $x \neq x_0$.)

Briefly, if $\alpha \sim \beta$ then $\alpha - \beta = o(\alpha)$ and vice versa.

Proof. Let $\alpha \sim \beta$. Then

$$\lim \frac{\alpha - \beta}{\alpha} = \lim \left(1 - \frac{\beta}{\alpha} \right) = 1 - \lim \frac{\beta}{\alpha} = 0$$

which exactly means that $\alpha - \beta = o(\alpha)$. The above relations also imply that $\alpha - \beta = o(\beta)$.

Now suppose it is given that α and β satisfy the condition $\alpha - \beta = o(\alpha)$. This means that $\lim \frac{\alpha - \beta}{\alpha} = 0$. Then $\lim \left(1 - \frac{\beta}{\alpha} \right) = 0$ whence $\lim \frac{\beta}{\alpha} = 1$, that is $\beta \sim \alpha$. The argument is similar if we start with the condition $\alpha - \beta = o(\beta)$.

Example. The infinitesimals $\alpha = \frac{x+1}{x^2}$ and $\beta = \frac{1}{x}$ are equivalent for $x \rightarrow \infty$ since their difference $\alpha - \beta = \frac{1}{x^2}$ is equal to $o(\beta)$. Indeed, we have $\lim_{x \rightarrow \infty} \frac{1}{x^2} : \frac{1}{x} = 0$. Therefore $\lim_{x \rightarrow \infty} \frac{\alpha}{\beta} = 1$, which can also be verified directly.

The properties of equivalent infinitesimals are often applied to finding various limits.

Let $\alpha \sim \beta$ and $\alpha_1 \sim \beta_1$. Then $\alpha \alpha_1 \sim \beta \beta_1$ and $\frac{\alpha}{\alpha_1} \sim \frac{\beta}{\beta_1}$ since $\frac{\alpha}{\alpha_1} \cdot \frac{\beta}{\beta_1} = \frac{\alpha}{\beta} \cdot \frac{\beta_1}{\alpha_1}$.

Furthermore, the identity $\frac{\alpha}{\alpha_1} = \frac{\alpha}{\beta} \cdot \frac{\beta}{\beta_1} \cdot \frac{\beta_1}{\alpha_1}$ shows that if the limit of the ratio $\frac{\beta}{\beta_1}$ exists the limit of the ratio $\frac{\alpha}{\alpha_1}$ exists as well, and they coincide:

$$\lim \frac{\alpha}{\alpha_1} = \lim \frac{\alpha}{\beta} \lim \frac{\beta}{\beta_1} \lim \frac{\beta_1}{\alpha_1} = \lim \frac{\beta}{\beta_1}$$

This can be stated briefly as *the limit of a ratio of infinitesimals does not change if any of the infinitesimals is replaced by an equivalent one.*

Examples. (1) We have $\lim_{x \rightarrow 0} \frac{\sin 2x}{\sin 5x} = \frac{2}{5}$ since $\sin 2x \sim 2x$ and $\sin 5x \sim 5x$.

(2) The infinitesimals $1 - \cos x = 2 \sin^2 \frac{x}{2}$ and $2 \cdot \frac{x}{2} \cdot \frac{x}{2} = \frac{x^2}{2}$ are equivalent as $x \rightarrow 0$, and therefore $\lim_{x \rightarrow 0} \frac{1 - \cos x}{x^2} = \frac{1}{2}$.

(3) We have $\lim_{x \rightarrow 0} \frac{\sin x + x^2}{x - x^2 + x^3} = 1$. Indeed, $\sin x + x^2 \sim x$ since $x^2 = o(x)$ and $x - x^2 + x^3 \sim x$ since $x^2 = o(x)$.

(c) Let $\alpha(x) \sim C\beta(x)$ for $x \rightarrow x_0$ (or for $x \rightarrow \infty$) where C is a nonzero constant. This means that

$$\lim_{x \rightarrow x_0} \frac{\alpha(x)}{\beta(x)} = C \neq 0$$

In this case $\alpha(x)$ and $\beta(x)$ are called *infinitesimals of the same order*. Equivalent infinitesimals are a particular case of infinitesimals of the same order (for $C=1$). As an example, we can take the infinitesimals $\sin 2x$ and $\sin 5x$ which are of the same order for $x \rightarrow 0$ (see Example 1).

II. Orders of infinitesimals. Let α and β be infinitesimals and k a positive number. If

$$\alpha \sim C\beta^k \text{ where } C \neq 0 \text{ is a constant}$$

α is said to be an *infinitesimal of the k th order relative to the infinitesimal β* .

This terminology is coherent with the fact that if x is an infinitesimal then x^k is an infinitesimal of the k th order relative to the infinitesimal x which thus plays the role of a "standard" infinitesimal magnitude.

The solution of Example 2 clearly indicates that the infinitesimal $1 - \cos x$ is of the second order relative to x as $x \rightarrow 0$.

It is evident that if $k > 1$ then α is an infinitesimal of higher order than β , if $k = 1$ they are of the same order and if $k < 1$ then, conversely, β is of higher order than α .

III. Some examples of equivalent infinitesimals connected with the number e . Natural logarithms. The function $\log_a(1+x)$ is an infinitesimal as $x \rightarrow 0$ since its limit is equal to $\log_a 1 = 0$. To compare it with x we must determine the limit of their ratio. To this end, we perform the following transformations:

$$\lim_{x \rightarrow 0} \frac{\log_a(1+x)}{x} = \lim_{x \rightarrow 0} \left[\frac{1}{x} \log_a(1+x) \right] = \lim_{x \rightarrow 0} \log_a(1+x)^{\frac{1}{x}}$$

By formula (**) (Sec. 34), the symbols of the logarithm and of the limit can be interchanged if the expression under the sign of logarithm possesses a limit as $x \rightarrow 0$. Let us prove the latter.

If we put $z = \frac{1}{x}$ then $z \rightarrow \infty$ as $x \rightarrow 0$, and therefore

$$\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = \lim_{z \rightarrow \infty} \left(1 + \frac{1}{z} \right)^z = e$$

(see Sec. 31). Now we conclude that

$$\lim_{x \rightarrow 0} \frac{\log_a(1+x)}{x} = \log_a \left[\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} \right] = \log_a e$$

We have thus shown that the infinitesimals $\log_a(1+x)$ and x are of the same order as $x \rightarrow 0$. If we put $a=e$ then $\log_e(1+x) = \ln(1+x)$ and $\log_e e = \ln e = 1$. The latter equality implies that $\ln(1+x)$ and x are equivalent infinitesimals as $x \rightarrow 0$:

$$\ln(1+x) \sim x \text{ for } x \rightarrow 0 \quad (*)$$

Now let us take the function $a^x - 1$ which is also an infinitesimal as $x \rightarrow 0$ and compare it with x . To this end we put

$$a^x - 1 = u$$

then $u \rightarrow 0$ as $x \rightarrow 0$. Taking the logarithms to the base a of both members of the equality $a^x = 1 + u$ we obtain $x = \log_a(1+u)$. Hence, $x \rightarrow 0$ as $u \rightarrow 0$, and

$$\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \lim_{u \rightarrow 0} \frac{u}{\log_a(1+u)} = \lim_{u \rightarrow 0} \frac{1}{\frac{\log_a(1+u)}{u}}$$

The limit of the function in the denominator is already known: it is equal to $\log_a e$. Therefore, $\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \frac{1}{\log_a e} = \log_e a = \ln a$, that is

$$a^x - 1 \sim x \ln a \text{ for } x \rightarrow 0 \quad (**)$$

Thus, we have proved that the infinitesimals $a^x - 1$ and x are of the same order for $x \rightarrow 0$. Putting $a=e$ we conclude that $e^x - 1$ and x are equivalent infinitesimals:

$$e^x - 1 \sim x \text{ for } x \rightarrow 0 \quad (***)$$

Relations (*) and (***) indicate that it is convenient to choose the number e as the base for the logarithmic and exponential functions. In this case $\ln(1+x)$ and $e^x - 1$ can be replaced by x for small values of $|x|$ with an arbitrarily small relative error as $|x| \rightarrow 0$:

$$\ln(1+x) \approx x, \quad e^x - 1 \approx x$$

This property simplifies many calculations and is the reason why in mathematical analysis the *exponential function* e^x and the *natural logarithm* $\ln x$ are so frequently used.

Logarithms of numbers to different bases being proportional (see Sec. 20), the passage from decimal logarithms to natural

ones and vice versa is readily carried out. The natural and decimal logarithms are connected by the relations

$$\log x = M \ln x \text{ and } \ln x = \frac{1}{M} \log x$$

where

$$M = \log e = \frac{1}{\ln 10} \approx 0.434294 \text{ and } \frac{1}{M} \approx 2.302585$$

In practical calculations we usually use decimal logarithms. They have no theoretical advantages over any other logarithms but are more convenient for computations.

37. Comparison of Infinitely Large Magnitudes.

1. The comparison of infinitely large magnitudes is performed by analogy with the comparison of infinitesimals but involves some changes in the terminology.

Let $u(x)$ and $v(x)$ be functions tending to infinity as $x \rightarrow x_0$ (or as $x \rightarrow \infty$).

If $\lim_{x \rightarrow x_0} \frac{u(x)}{v(x)} = 0$ the symbolic relation $u(x) = o(v(x))$ is retained, but in this case we say that the *denominator* $v(x)$ is an *infinitely large magnitude of higher order than the numerator* $u(x)$ (or that $u(x)$ is of *lower order* than $v(x)$).

According to this change of terminology, the power functions $x, x^2, x^3, \dots$ form a sequence of infinitely large magnitudes as $x \rightarrow \infty$ with increasing orders: $x = o(x^2)$, $x^2 = o(x^3)$, etc.

If $\lim_{x \rightarrow x_0} \frac{u(x)}{v(x)} = 1$ we retain the notation $u(x) \sim v(x)$ but say that $u(x)$ and $v(x)$ are *asymptotically equivalent*.

As before, the asymptotic relation $u \sim v$ is equivalent to $u - v = o(u)$ (or to $u - v = o(v)$).

In the determination of the limit of a quotient of two infinitely large magnitudes the rule allowing us to replace any of these magnitudes by an (asymptotically) equivalent expression is also retained:

$$\text{if } u \sim v \text{ and } u_1 \sim v_1 \text{ then } \lim \frac{u}{v} = \lim \frac{u_1}{v_1}$$

on condition that the limits involved exist.

II. If $u(x)$ and $v(x)$ are infinitely large magnitudes the asymptotic relation $u \sim v$ does not at all mean that the absolute value of their difference $|u(x) - v(x)|$ is an infinitesimal; in other words, the absolute error appearing when one of the functions is replaced by an asymptotically equivalent function is not necessarily small. But, at the same time, the relative error $\left| \frac{u-v}{u} \right|$ becomes an infinitesimal, and it is this quantity that characterizes the accuracy of the substitution of v for u (see Sec. 6).

Hence, every asymptotical relation $u \sim v$ can be interpreted as an approximate formula $u(x) \approx v(x)$ allowing us to replace one of the functions by another function (which is often simpler) with a high relative accuracy.

For instance, let $u = x^2 + \sqrt{x} \arctan x$ and $v = x^2$. Here we have $u - v = \sqrt{x} \arctan x = o(x^2)$ as $x \rightarrow +\infty$ since $\lim_{x \rightarrow \infty} \frac{\sqrt{x} \arctan x}{x^2} = \lim_{x \rightarrow \infty} \frac{\arctan x}{x \sqrt{x}} = 0$ (because the numerator of the last fraction is a bounded function and the denominator an infinitely large magnitude). Consequently, the approximate equality $x^2 + \sqrt{x} \arctan x \approx x^2$ has a high relative accuracy for large values of x .

There is a remarkable asymptotic relation known as *Stirling's formula* which we give here without proof:

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \text{ for } n \rightarrow \infty$$

This formula makes it possible to compute $n!$ for large values of n . The practical application of Stirling's formula is based on computing its right-hand member by means of logarithms (the common logarithm of the right member is equal to $\frac{1}{2} \log 2\pi + \left(n + \frac{1}{2}\right) \log n - n \log e$).

Stirling's formula is widely used in the *theory of probability*.

III. In conclusion we note that the symbolic relation $\varphi(x) = o(f(x))$ for $x \rightarrow x_0$ (or for $x \rightarrow \infty$) is used in all cases when $\lim_{x \rightarrow x_0} \frac{\varphi(x)}{f(x)} = 0$ irrespective of the properties of the functions $\varphi(x)$ and $f(x)$ (provided that the denominator does not vanish). For instance, the relation $\alpha(x) = o(1)$ means that $\alpha(x)$ is an infinitesimal.

Similarly, the symbolic relation $\varphi(x) \sim f(x)$ always means that $\lim_{x \rightarrow x_0} \frac{\varphi(x)}{f(x)} = 1$ and always implies that $\varphi(x) - f(x) = o(\varphi(x))$ (or $\varphi(x) - f(x) = o(f(x))$).

Sometimes we also use the symbol "O": the relation

$$f(x) = O(g(x)) \text{ for } x \rightarrow x_0$$

(read " $f(x)$ is big O of $g(x)$ for $x \rightarrow x_0$ ") means that, for all $x \neq x_0$ belonging to the neighbourhood of the point x_0 , the inequality

$$|f(x)| \leq C |g(x)|$$

¹ J. Stirling (1692-1770), an English mathematician.

is fulfilled where $C > 0$ is a constant. The relation $f(x) = O(1)$ as $x \rightarrow x_0$ indicates that $f(x)$ is a bounded function in the neighbourhood of the point x_0 .

The relation $f(x) = O(g(x))$ for $x \rightarrow \infty$ is understood in the analogous sense.

QUESTIONS

1. State the definition of the limit of a function $y = f(x)$ as $x \rightarrow x_0$. Express the definition by means of inequalities and explain its geometrical interpretation.

2. Give examples of functions $y = f(x)$ possessing a limit as $x \rightarrow x_0$ and having no limit as $x \rightarrow x_0$.

3. What do we call the limit of a function $y = f(x)$ as $x \rightarrow +\infty$ ($x \rightarrow -\infty$)? Express the definition by means of inequalities and give its geometrical interpretation.

4. Give examples of functions $y = f(x)$ possessing limits as $x \rightarrow +\infty$ (as $x \rightarrow -\infty$) and having no limits as $x \rightarrow +\infty$ (as $x \rightarrow -\infty$).

5. What is the limit of a sequence? Give examples of a sequence possessing a limit and of one having no limit.

6. What is meant by saying that the function $y = f(x)$ tends to infinity as $x \rightarrow x_0$ (as $x \rightarrow \infty$)? Express the definition by means of inequalities. Give geometrical illustrations.

7. What do the relations

$$\begin{aligned} \lim_{x \rightarrow x_0} f(x) = +\infty, \quad \lim_{x \rightarrow +\infty} f(x) = +\infty, \quad \lim_{x \rightarrow -\infty} f(x) = +\infty \\ \lim_{x \rightarrow -\infty} f(x) = +\infty, \quad \lim_{x \rightarrow x_0} f(x) = -\infty, \quad \lim_{x \rightarrow +\infty} f(x) = -\infty \quad \text{and} \\ \lim_{x \rightarrow -\infty} f(x) = -\infty, \quad \lim_{x \rightarrow \infty} f(x) = -\infty \end{aligned}$$

mean? Give verbal explanation and also express the corresponding definitions by means of inequalities. Give geometrical interpretation.

8. Give illustrative examples of functions tending to infinity for various types of the limiting behaviour of the argument.

9. When do we say that a function is bounded on an interval or is bounded for $x \rightarrow x_0$ or for $x \rightarrow \infty$?

10. Give an example of a function which is unbounded but is not infinitely large.

11. What function $y = f(x)$ is said to be an infinitesimal as $x \rightarrow x_0$ (as $x \rightarrow \infty$)? Express the definitions by means of inequalities and give their geometrical interpretation.

12. What is the relationship between an infinitely large magnitude and an infinitesimal? Prove the corresponding theorem.

13. How is the fact that a function possesses a limit is connected with the notion of an infinitesimal? Prove the corresponding direct and reciprocal theorems.

14. State and prove the rules for arithmetical operations on magnitudes possessing limits.

15. State and prove the test for the existence of the limit of a function whose values lie between the values of two functions having one and the same limit.

16. Prove that $\lim_{\alpha \rightarrow 0} \frac{\sin \alpha}{\alpha} = 1$.

17. State the test for the existence of the limit of a monotone bounded sequence. Explain its geometrical meaning.

18. Prove that the limit of the sequence $\left(1 + \frac{1}{n}\right)^n$, as $n \rightarrow \infty$, exists.

19. State the definition of the continuity of a function $y = f(x)$ at a point x_0 and illustrate it geometrically.

20. What is a point of discontinuity of a function?

21. Give examples of various types of a point of discontinuity.

22. State the rule for passing to the limit for a continuous function.

23. State and prove the theorems on arithmetical operations on continuous functions.

24. State and prove the theorem on the continuity of a composite function formed of continuous functions.

25. What can you say about the intervals of continuity of an elementary function? What points can be the points of discontinuity of an elementary function?

26. State the properties of functions continuous in a closed interval. Demonstrate them geometrically.

27. What does "to compare two infinitesimals" mean? When is one of them said to be of higher order than the other?

28. When are two infinitesimals called equivalent? State the necessary and sufficient condition for the equivalence.

29. Give examples of equivalent infinitesimals.

30. Give the numerical values of the limits $\lim_{x \rightarrow 0} \frac{a^x - 1}{x}$ and $\lim_{x \rightarrow 0} \frac{\log_a(1+x)}{x}$.

31. What is the system of natural logarithms? How do we pass from decimal logarithms to natural ones and vice versa?

32. How do we compare two infinitely large magnitudes? When one of them is said to be of higher order than the other?

33. What do we call asymptotically equivalent infinitely large magnitudes? Give examples of such magnitudes.

Chapter III

DERIVATIVE AND DIFFERENTIAL. DIFFERENTIAL CALCULUS

§ 1. Derivative

38. Some Physical Problems.

I. Velocity of rectilinear motion. Let a point move in a straight line which is taken as the number scale and let the law of variation of the coordinate s of the moving point as function of time t be known:

$$s = F(t)$$

The equation $s = F(t)$ is called the *equation (the law) of motion* and the curve in the ts -plane specified by it is called the *graph of motion*¹.

During time interval Δt from time moment t to $t + \Delta t$ the coordinate of the point gains the increment

$$\Delta s = F(t + \Delta t) - F(t)$$

(if the function $F(t)$ is monotone in the interval $[t, t + \Delta t]$ the number $|\Delta s|$ expresses the path length travelled during time Δt).

If the motion is uniform, that is, s is a linear function of t of the form $s = v_0 t + s_0$, we have $\Delta s = v_0 \Delta t$ (see Sec. 15), and $\frac{\Delta s}{\Delta t} = v_0$ is a constant *velocity of the rectilinear motion of the point*.

But if the motion is nonuniform the ratio $\frac{\Delta s}{\Delta t}$ depends both on t and on Δt . It is then called the *average (mean) velocity* corresponding to time interval from t to $t + \Delta t$. Denoting it v_{av} we can write

$$v_{av} = \frac{\Delta s}{\Delta t}$$

For a given time interval and a definite path length the point having a fixed average velocity can move in various ways. This is clearly seen in Fig. 49 which shows various lines (AC_1B , AC_2B , AC_3B)

¹ The graph of motion is a curve in the coordinate system Ots (Fig. 49) and must not be confused with the *trajectory of motion*.

and AC_3B) connecting two points A and B in the ts -plane. These lines are the graphs of motion to which one and the same average velocity $\frac{\Delta s}{\Delta t}$ corresponds although the character of motion varies. In particular, the rectilinear segment ACB joining the point A to the point B is the graph of the uniform motion with velocity v_{av} during time interval $[t, t + \Delta t]$. Thus, the *average velocity* $\frac{\Delta s}{\Delta t}$ is the velocity of the uniform motion in which the point covers the distance Δs during the time interval $[t, t + \Delta t]$.

Now let us retain the value of t and decrease Δt . The average velocity v_{av} corresponding to the new time interval $[t, t + \Delta t_1]$ ($\Delta t_1 < \Delta t$) lying inside the original interval $[t, t + \Delta t]$ may of course differ from the average velocity for the whole interval $[t, t + \Delta t]$. This means that the average velocity does not characterize a nonuniform motion completely since it depends on the time interval it is related to. The shorter the time interval $[t, t + \Delta t]$ (i.e. the smaller Δt), the better the average velocity characterizes the motion. Therefore it appears natural to make Δt tend to zero. If, in this process, the average velocity v_{av} possesses a limit, it is natural to consider this limit as the ("true") velocity at the given time t .

Definition. The limit v of the average velocity v_{av} corresponding to time interval $[t, t + \Delta t]$, as $\Delta t \rightarrow 0$, is called the (*instantaneous*) *velocity of rectilinear motion at the given moment t* :

$$v = \lim_{\Delta t \rightarrow 0} v_{av} = \lim_{\Delta t \rightarrow 0} \frac{\Delta s}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t}$$

Example. Let us consider the law of free fall:

$$s = \frac{1}{2} g t^2$$

Computing the average velocity for a time interval $(t, t + \Delta t)$ we find

$$v_{av} = \frac{\frac{1}{2} g (t + \Delta t)^2 - \frac{1}{2} g t^2}{\Delta t} = \frac{1}{2} g \frac{2t\Delta t + (\Delta t)^2}{\Delta t} = \frac{1}{2} g (2t + \Delta t)$$

Now we determine the velocity at the moment t :

$$v = \lim_{\Delta t \rightarrow 0} \frac{1}{2} g (2t + \Delta t) = g t$$

Thus, the velocity of free fall is directly proportional to the time of motion.

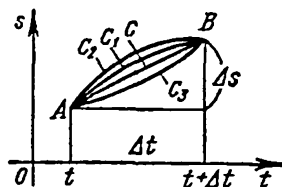


Fig. 49

II. Density. We now proceed to discuss the notion of a *mass density* playing an important role in problems of physics. The scheme of the argument is analogous to the above.

Suppose we are given a material line (a wire, for instance). Let us reckon the path length s along the line and the mass m of the corresponding portion of the wire in the direction from one of its ends to the other. To each value of s there corresponds a definite value of m which thus is a function of s :

$$m = \Phi(s)$$

We say that the mass of the line is distributed along it *uniformly* or that the material line is *homogeneous* if the masses of any two portions of the line having equal lengths coincide. In this case m is a linear function of s (more precisely, m is directly proportional to s): $m = \delta_0 s$ where δ_0 is the constant of variation. We have $\Delta m = \delta_0 \Delta s$, and the ratio $\frac{\Delta m}{\Delta s}$ (which is equal to δ_0 in this case) shows how many mass units are carried by the portion of the line of unit lengths. This constant ratio is called the (*linear*) *mass density* of the homogeneous material line.

Now suppose that the distribution of mass is not uniform. Take the portion of the line from s to $s + \Delta s$; its mass Δm is given by the formula

$$\Delta m = \Phi(s + \Delta s) - \Phi(s)$$

The ratio $\frac{\Delta m}{\Delta s}$ of the mass to the length is no longer constant and depends on s and on Δs . It is called the *mean (average) mass density* of the line (of the wire) corresponding to the part $[s, s + \Delta s]$ of that line:

$$\delta_{av} = \frac{\Delta m}{\Delta s}$$

A mass Δm can be distributed along the interval $[s, s + \Delta s]$ in various ways and, in particular, uniformly. Hence, the mean density $\frac{\Delta m}{\Delta s}$ shows with what constant density the mass should be distributed *uniformly* along the interval $[s, s + \Delta s]$ so that the total mass remains the same. Therefore the mean density does not provide a complete characteristic of the mass distribution and only describes it "in the mean". To avoid this ambiguity we introduce the notion of the *linear density at a given point* (for a given value of s). The latter notion is defined on the basis of the fact that the smaller the portion Δs of the line, the more accurate is the description of the mass distribution in the vicinity of the point s given by the average density. Therefore we state that the *linear density* δ of a material line at a given point s is,

by definition, the limit of the average density δ_{av} corresponding to the interval $[s, s + \Delta s]$ as Δs tends to zero:

$$\delta = \lim_{\Delta s \rightarrow 0} \delta_{av} = \lim_{\Delta s \rightarrow 0} \frac{\Delta m}{\Delta s} = \lim_{\Delta s \rightarrow 0} \frac{\Phi(s + \Delta s) - \Phi(s)}{\Delta s}$$

III. Heat capacity. Reaction rate. Here we briefly discuss two more notions whose definition involves the same argument as in the cases of velocity and density, namely the notions of *heat capacity* and of *reaction rate* (in chemistry).

The *heat capacity* c of a physical body for temperature τ is the limit of the average heat capacity c_{av} corresponding to the interval $[\tau, \tau + \Delta \tau]$ as $\Delta \tau$ tends to zero:

$$c = \lim_{\Delta \tau \rightarrow 0} c_{av} = \lim_{\Delta \tau \rightarrow 0} \frac{\Delta q}{\Delta \tau} = \lim_{\Delta \tau \rightarrow 0} \frac{\Psi(\tau + \Delta \tau) - \Psi(\tau)}{\Delta \tau}$$

where $q = \Psi(\tau)$ is the quantity of heat absorbed by the body as it is heated to the temperature τ .

The *reaction rate* γ for a given amount of the substance involved at a given time moment t is the limit of the average reaction rate γ_{av} corresponding to the time interval $[t, t + \Delta t]$ as Δt tends to zero:

$$\gamma = \lim_{\Delta t \rightarrow 0} \gamma_{av} = \lim_{\Delta t \rightarrow 0} \frac{\Delta m}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t}$$

where $m = F(t)$ is the mass of the substance entering into the reaction during time t .

39. The Rate of Change of a Function at a Point. The Derivative of a Function. The Derivative of the Power Function.

I. The rate of change of a function. The four notions (velocity of motion, density, heat capacity and reaction rate) discussed in Sec. 38 are of different physical nature but, from the mathematical point of view, they give the same characteristic of the corresponding function. Namely, they all are special cases of the so-called *rate of change of a function* which is defined in an analogous manner by means of the concept of limit.

Let us proceed to the general discussion of the notion of the rate of change of a function $y = f(x)$ abstracting from the physical meaning of the variables x and y .

To begin with, let us take a linear function

$$y = f(x) = ax + b$$

If the independent variable x is given an increment Δx the function y receives the corresponding increment $\Delta y = a \Delta x$. Here the ratio $\frac{\Delta y}{\Delta x} = a$ is a *constant* independent both of the value of x for which the value of the function is taken and of the given value

of Δx . This ratio is called the *rate of change* of the linear function.

But if the function $y = f(x)$ is nonlinear the ratio

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

depends both on x and on Δx . It only characterizes "in the mean" the rate of variation of the function as the independent variable ranges from x to $x + \Delta x$; it is equal to the rate of change of a *linear* function which receives the same increment Δy for the same value of Δx .

The ratio $\frac{\Delta y}{\Delta x}$ is called the *average (mean) rate of change* v_{av} of the function $y = f(x)$ in the interval $[x, x + \Delta x]$.

It is clear that the smaller the interval under consideration, the more precise the characterization of the variation of the function given by the quantity v_{av} . Therefore it appears natural to consider the process in which Δx is made to tend to zero. If, in this process, the limit of the average rate exists it is taken as the measure of the rate of change of the function for the given value of x and is called the *rate of change of the function*.

Definition. The *rate of change v of a function $f(x)$ at a given point x* is the limit of the average rate of change of that function in the interval $[x, x + \Delta x]$ as Δx tends to zero (on condition that this limit exists):

$$v = \lim_{\Delta x \rightarrow 0} v_{av} = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

II. The derivative of a function. The rate of change of a function $y = f(x)$ is defined by means of the following sequence of operations:

(1) given an arbitrary increment Δx of the argument, the corresponding increment of the function is found:

$$\Delta y = f(x + \Delta x) - f(x)$$

(2) the ratio $\frac{\Delta y}{\Delta x}$ is formed;

(3) the limit of this ratio (provided that it exists) as Δx tends to zero is found.

As was mentioned, if the given function $f(x)$ is nonlinear the ratio $\frac{\Delta y}{\Delta x}$ depends both on x and on Δx . When Δx is made to tend to zero the resultant limit of the ratio $\frac{\Delta y}{\Delta x}$ depends *solely on the chosen value of x* and thus is a function of x . (If the function $f(x)$ is linear this limit is independent of x and is a constant magnitude.)

The limit thus obtained is called *the derivative of the function $f(x)$ and is a new function of x denoted by $f'(x)$* (read "f dash x").

Definition. The limit of the ratio of the increment of a given function $y=f(x)$ to the increment of the independent variable as the latter tends to zero is called the *derivative* of that function (provided that this limit exists):

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

The *particular value* of the derivative $f'(x)$ at a given point x_0 is usually denoted by $f'(x_0)$ or $y'_{x=x_0}$.

Suppose that the derivative $f'(x)$ at a point x exists; this means that there exists a finite limit $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x)$. Then the ratio $\frac{\Delta y}{\Delta x}$ itself can be represented in the form $\frac{\Delta y}{\Delta x} = f'(x) + \varepsilon$ where ε is an infinitesimal as $\Delta x \rightarrow 0$, whence

$$\Delta y = f'(x) \Delta x + \varepsilon \Delta x$$

Thus, if $\Delta x \rightarrow 0$ then $\Delta y \rightarrow 0$, and therefore the function $f(x)$ is continuous at the point x . Consequently, *if a function $f(x)$ has a derivative at a point x it is continuous at that point.*

It turns out that the converse assertion is not true: there are functions continuous at a point and having no derivative at that point. Examples of such functions will be discussed later on.

Now, using the notion of the derivative we can say that

(1) the velocity of a rectilinear motion of a point is the derivative of the function $s=F(t)$ with respect to t (that is, the derivative of the *coordinate* with respect to *time*);

(2) the linear (mass) density is the derivative of the function $m=\Phi(s)$ with respect to s (that is the derivative of the *mass* with respect to the *length*);

(3) the heat capacity is the derivative of the function $q=\Psi(\tau)$ with respect to τ (that is, the derivative of the *amount of heat* with respect to *temperature*);

(4) the reaction rate is the derivative of the function $\gamma=F(t)$ with respect to t (the derivative of the *mass of substance* with respect to *time*).

Generally, *the rate of change of a function $y=f(x)$ is the derivative of y with respect to x .*

The notion of the derivative is one of the fundamental concepts of mathematical analysis. I. Newton¹ was one of the founders of mathematical analysis; he introduced the notion of the deriva-

¹ I. Newton (1642-1727), the great English scientist. He contributed many fundamental results to mathematics, physics and astronomy.

tive when studying the problem of determination of the velocity of motion.

111. The derivative of the power function. Let us find the derivatives of some simple functions. To begin with, take $y=x$. Then

$$y' = (x)' = \lim_{\Delta x \rightarrow 0} \frac{(x+\Delta x) - x}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{\Delta x}{\Delta x} = 1$$

that is the derivative $(x)'$ is a constant equal to 1. This is obvious since $y=x$ is a linear function and its rate of change must be constant.

If $y=x^2$ then

$$y' = (x^2)' = \lim_{\Delta x \rightarrow 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} = \lim_{\Delta x \rightarrow 0} (2x + \Delta x) = 2x$$

If $y=x^3$ then

$$y' = (x^3)' = \lim_{\Delta x \rightarrow 0} \frac{(x+\Delta x)^3 - x^3}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3}{\Delta x} = 3x^2$$

Now we can observe the general feature of the structure of the derivatives of the power function $y=x^n$ for $n=1, 2, 3$. Let us prove that, generally, the derivative of $y=x^n$ for any positive integral exponent n is equal to nx^{n-1} .

To this end, we write

$$\frac{\Delta y}{\Delta x} = \frac{(x+\Delta x)^n - x^n}{\Delta x}$$

The expression in the numerator is transformed with the help of Newton's binomial formula:

$$(x+\Delta x)^n - x^n = x^n + nx^{n-1}\Delta x + \frac{n(n-1)}{2!}x^{n-2}\Delta x^2 + \dots + \Delta x^n - x^n$$

Hence,

$$\frac{\Delta y}{\Delta x} = nx^{n-1} + \frac{n(n-1)}{2!}x^{n-2}\Delta x + \dots + \Delta x^{n-1}$$

The right member of the last equality is a sum of n terms; the first term is independent of Δx and the others tend to zero as $\Delta x \rightarrow 0$. Therefore,

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = nx^{n-1}$$

Thus, for every positive integral exponent n , the power function $y=x^n$ has the derivative nx^{n-1} :

$$(x^n)' = nx^{n-1}$$

Putting $n=1, 2, 3$ in this general formula we receive the special results obtained above.

Later on (in Sec. 45) we shall show that this formula applies to any exponent n as well. For instance,

$$(V\bar{x})' = \left(x^{\frac{1}{2}}\right)' = \frac{1}{2} x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}$$

Let us consider separately the derivative of a constant magnitude

$$y = C$$

The value of the function being invariable as the independent variable changes, we have $\Delta y = 0$ and $\frac{\Delta y}{\Delta x} = \frac{0}{\Delta x} = 0$. Consequently,

$$(C)' = 0$$

that is, the derivative of a constant is equal to zero.

40. Geometrical Meaning of the Derivative. The derivative of a function $y=f(x)$ admits of a simple and visual geometrical interpretation closely related to the notion of the *tangent line* (or, briefly, the *tangent*) to a curve.

Definition. The *tangent line* M_0T (Fig. 50) to a curve AB at its point M_0 is the straight line occupying the limiting position of the line passing through the point M_0 and another point M of the curve as M moves along the curve to coincide with the given point M_0 .

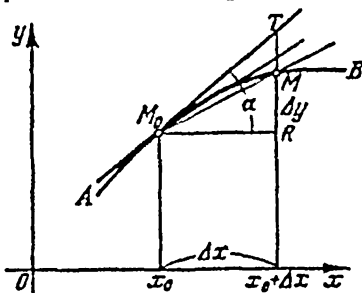


Fig. 50

As the line passing through M_0 and M tends to its limiting position the angle TM_0M tends to zero together with the chord M_0M .

The geometrical meaning of the derivative is implied by the following proposition:

Theorem. If the value of the derivative of a function $y=f(x)$ at a point $x=x_0$ is equal to $f'(x_0)$ ¹ the straight line with slope $f'(x_0)$ passing through the point $M_0(x_0, y_0)$ ($y_0=f(x_0)$) is the tangent to the graph of the function $y=f(x)$ at the point M_0 .

Proof. Let us draw through the point M_0 the straight line M_0T with slope $f'(x_0)$ (see Fig. 50). Then $f'(x_0) = \tan \alpha$ where α is the angle of inclination of the line M_0T to the axis of abscissas.

¹ It should be stressed that the value of the derivative $f'(x_0)$ is found in the following way: we first compute the derivative $f'(x)$ as function of x and then substitute into it the given concrete value x_0 of the independent variable for x . It would be absurd to begin with the value $f(x_0)$ of the given function at the point x_0 and then differentiate it since $f(x_0)$ is a constant and its derivative, equal to zero, has nothing in common with the value of the derivative $f'(x_0)$ we are interested in!

Now we give $x=x_0$ an increment Δx , take the point M of the graph of the function corresponding to the value of the argument $x=x_0+\Delta x$ and draw the secant M_0M . The slope of the secant is equal to $\frac{RM}{M_0R}=\frac{\Delta y}{\Delta x}$ where $\Delta y=f(x_0+\Delta x)-f(x_0)$.

Now let $\Delta x \rightarrow 0$; then the point M moves along the arc AB toward the point M_0 . In this process the secant M_0M turns about the point M_0 and, according to the conditions of the theorem, its slope tends to the limit

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x_0) \quad (*)$$

which is equal to the slope of the line M_0T . Applying the formula for the acute angle between two straight lines (M_0T and M_0M) we obtain

$$\tan \angle TM_0M = \frac{\left| f'(x_0) - \frac{\Delta y}{\Delta x} \right|}{1 + f'(x_0) \frac{\Delta y}{\Delta x}}$$

By equality (*), if $\Delta x \rightarrow 0$ the numerator of the fraction tends to zero while the denominator to the number $1 + [f'(x_0)]^2 \neq 0$. Therefore $\tan \angle TM_0M$ tends to zero, and hence $\angle TM_0M$ also tends to zero¹. We have thus proved that the straight line M_0T is the tangent. This theorem establishing the geometrical meaning of the derivative can be stated briefly as follows:

The value of the derivative $f'(x_0)$ is equal to the slope of the tangent line to the graph of the function $y=f(x)$ at the point with abscissa x_0 .

Remark: G.W. Leibniz² who, at the same time as I. Newton, introduced the notion of the derivative proceeded from the problem of constructing the tangent to a curve.

§ 2. Differentiating Functions

41. Differentiation and Arithmetical Operations. The process of finding the derivative of a function is referred to as the *differentiation* of the function (the explanation of this terminology will be given in Sec. 50). "To differentiate a function $f(x)$ " means to

¹ The angle $\angle TM_0M = \varphi$ being acute, the condition $\tan \varphi = z \rightarrow 0$ implies $\varphi = \arctan z \rightarrow 0$.

² G.W. Leibniz (1645-1716) was the great German philosopher and mathematician who contributed many outstanding results to various divisions of science (mathematics, physics, philosophy, history, etc.). At the same time as I. Newton he discovered differential and integral calculus and founded mathematical analysis.

compute its derivative $f'(x)$. To this end, we must determine the limit

$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

The direct evaluation of such a limit is most often connected with lengthy and complicated calculations. But it turns out that for the basic elementary functions it is possible to derive general formulas expressing their derivatives analytically (remember that by now we only know the formula for the derivative of the power function $y = x^n$). Therefore, since the rules for differentiating composite functions and those resulting from arithmetical operations on functions are also readily established, we can always find analytically the derivative of *any elementary function* without resorting to the computation of the limit indicated above.

We shall begin with establishing rules for differentiating the results of arithmetical operations. The constituent functions of which sums, products and quotients are formed will be supposed to possess derivatives for the given values of the independent variable. The proofs given below imply that the functions resulting from these operations, that is, sums, products and quotients, also have derivatives for the same values of the independent variable. Bearing this remark in mind we shall not stipulate the existence of the derivative each time and shall only confine ourselves to the question of finding the derivative of the resulting function when the derivatives of the constituent functions are known.

Theorem I. The derivative of a sum of a finite number of functions is equal to the sum of the derivatives of the summands.

Proof. Let y be a function represented as the sum of given functions $u, v, \dots, w$ of the same independent variable x :

$$y = u + v + \dots + w$$

We have to prove that

$$y' = (u + v + \dots + w)' = u' + v' + \dots + w'$$

If the independent variable x is given an increment Δx the functions $u, v, \dots, w$ gain the corresponding increments $\Delta u, \Delta v, \dots, \Delta w$ and the function y the increment Δy . It is clear that

$$y + \Delta y = (u + \Delta u) + (v + \Delta v) + \dots + (w + \Delta w)$$

Subtracting from this relation the function y we obtain

$$\Delta y = \Delta u + \Delta v + \dots + \Delta w$$

On dividing the latter relation by Δx we receive

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x} + \dots + \frac{\Delta w}{\Delta x}$$

Now, passing to the limit as $\Delta x \rightarrow 0$ and taking advantage of the theorem on the limit of a sum (Sec. 29) we derive

$$\begin{aligned}\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} &= \lim_{\Delta x \rightarrow 0} \left(\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x} + \dots + \frac{\Delta w}{\Delta x} \right) = \\ &= \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} + \dots + \lim_{\Delta x \rightarrow 0} \frac{\Delta w}{\Delta x}\end{aligned}$$

We have

$$\begin{aligned}\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} &= y', \quad \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = u', \quad \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} = v', \quad \dots \\ &\dots, \quad \lim_{\Delta x \rightarrow 0} \frac{\Delta w}{\Delta x} = w'\end{aligned}$$

and thus arrive at the desired result.

Examples. (1) Consider the sum $y = x^3 - x^2 + x - 1$. Applying the above rule we find

$$y' = (x^3)' - (x^2)' + (x)' - (1)' = 3x^2 - 2x + 1$$

(2) Taking the function $y = x + \frac{1}{x}$ and performing the differentiation we obtain

$$y' = (x)' + (x^{-1})' = 1 - x^{-2} = 1 - \frac{1}{x^2}$$

Theorem II. The derivative of the product of two functions is equal to the sum of the product of the derivative of the first function by the second function and the product of the derivative of the second function by the first function.

Proof. Let a function y be represented as a product of two functions:

$$y = uv$$

We shall prove that

$$y' = (uv)' = u'v + v'u$$

Giving the independent variable an increment Δx and taking the corresponding increments Δu , Δv , Δy of the functions u , v , y we can write

$$y = \Delta y = (u + \Delta u)(v + \Delta v)$$

It follows that

$$\Delta y = (y + \Delta y) - y = (u + \Delta u)(v + \Delta v) - uv = \Delta u \cdot v + u \Delta v + \Delta u \Delta v$$

and

$$\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} v + u \frac{\Delta v}{\Delta x} + \frac{\Delta u}{\Delta x} \Delta v$$

Now, applying the rules for passing to the limit and taking into account that if $\Delta x \rightarrow 0$ then Δv as the increment of a con-

tinuous function also tends to zero, we obtain

$$\begin{aligned}\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} &= \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} v + u \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} + \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} \lim_{\Delta x \rightarrow 0} \Delta v = \\ &= u'v + uv' + u' \cdot 0 = u'v + uv'\end{aligned}$$

which is what we set out to prove.

If one of the factors is a constant magnitude, for instance, $v=C$ we receive

$$y' = (Cu)' = C'u + Cu' = Cu'$$

because the derivative of a constant is equal to zero. Hence, *a constant factor can be taken outside the differentiation sign.*

The differentiation rule for the product of two functions is readily extended to the product of any finite number of functions. For instance, if $y = uvw$ we write

$$\begin{aligned}y' = (uvw)' &= (uv)'w + uvw' = (u'v + uv')w + uvw' = \\ &= u'vw + uv'w + uvw'\end{aligned}$$

that is, the derivative of the product of three functions is equal to the sum of three summands each of which is the product of two of the given functions by the derivative of the remaining function.

Examples. (1) For the function $y = 5x^3 - 2x^2 + 3x - 1$, we have

$$y' = 5(x^3)' - 2(x^2)' + 3(x)' - (1)' = 15x^2 - 4x + 3$$

(2) Consider $y = (2x+3)(1-x)(x+2)$. The differentiation results in

$$\begin{aligned}y' &= (2x+3)'(1-x)(x+2) + (2x+3)(1-x)'(x+2) + \\ &\quad + (2x+3)(1-x)(x+2)' = \\ &= 2(1-x)(x+2) - (2x+3)(x+2) + (2x+3)(1-x)\end{aligned}$$

The same result is of course obtained if we open the parentheses in the given product and take the derivative of the resultant polynomial of the third degree.

Theorem III. The derivative of the quotient of two functions is equal to the fraction whose denominator is equal to the square of the divisor and the numerator is equal to the difference between the product of the derivative of the dividend by the divisor and the product of the dividend by the derivative of the divisor.

Proof. Let y be a function equal to the quotient of two functions u and v :

$$y = \frac{u}{v}$$

We must prove that

$$y' = \frac{u'v - uv'}{v^2}$$

Indeed, we have

$$\frac{\Delta y}{\Delta x} = \frac{\frac{u + \Delta u}{v + \Delta v} - \frac{u}{v}}{\Delta x} = \frac{\Delta u \cdot v - u \Delta v}{\Delta x \cdot v (v + \Delta v)} = \frac{\frac{\Delta u}{\Delta x} v - u \frac{\Delta v}{\Delta x}}{v (v + \Delta v)}$$

Passing to the limit as $\Delta x \rightarrow 0$ and taking into account that Δv as increment of a continuous function tends to zero together with Δx we receive

$$y' = \frac{v \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} - u \lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x}}{v \left(v + \lim_{\Delta x \rightarrow 0} \Delta v \right)} = \frac{u'v - uv'}{v^2}$$

which is what we had to prove. It is of course supposed here that $v \neq 0$ at the given point since, if otherwise, the function y is not defined at that point.

Examples. (1) Computing the derivative of the function $y = \frac{2-3x}{2+x}$ we find

$$y' = \frac{(2-3x)'(2+x) - (2-3x)(2+x)'}{(2+x)^2} = \frac{-3(2+x) - (2-3x)}{(2+x)^2} = -\frac{8}{(2+x)^2}$$

(2) Let $y = x^{-n}$ where n is a positive integer. We obtain

$$y' = \left(\frac{1}{x^n} \right)' = \frac{(1)'x^n - 1(x^n)'}{x^{2n}} = \frac{-nx^{n-1}}{x^{2n}} = -nx^{-n-1}$$

which confirms the rule for differentiating the power function in the case of a negative integral exponent.

(3) Let $y = \frac{1}{v}$ where v is a function of x . The derivative of the numerator being equal to zero, the formula for differentiating a quotient leads to

$$y' = -\frac{v'}{v^2}$$

It is advisable to memorize this result and its special case for $v = x$:

$$\left(\frac{1}{x} \right)' = -\frac{1}{x^2}$$

(4) Differentiate $y = \frac{u}{C}$ where C is a constant. It is unnecessary to differentiate this function as a quotient since it can be simply represented as the product of the constant factor $\frac{1}{C}$ by the function u :

$$y = \frac{1}{C} u$$

and hence

$$y' = \left(\frac{1}{C} u \right)' = \frac{1}{C} u' = \frac{u'}{C}$$

42. Differentiating Composite and Inverse Functions.

1. Differentiating composite function. The rule for differentiating function of a function is extremely important; it expresses the sought-for derivative in terms of the derivatives of the constituent functions the given function is formed of (on condition at these functions have derivatives).

Theorem. The derivative of a function of a function is equal the product of the derivative of the given function with respect the intermediate argument by the derivative of this argument with respect to the independent variable.

Proof. Let $y=f(u)$ and $u=\varphi(x)$. We have to prove that

$$y' = f'(u) u' = f'(u) \varphi'(x)$$

Let x receive an increment Δx . This results in an increment Δu the intermediate argument $u=\varphi(x)$ which, in its turn, generates an increment Δy of the magnitude y . To find y' we must compute $\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ as $\Delta x \rightarrow 0$. Let us represent the ratio $\frac{\Delta y}{\Delta x}$ in the form

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u} \frac{\Delta u}{\Delta x}$$

According to the rules for passing to the limit in a product we can write

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x}$$

where $\Delta u \rightarrow 0$, as $\Delta x \rightarrow 0$, since $u=\varphi(x)$ is a continuous function). Since

$$\lim_{\Delta u \rightarrow 0} \frac{\Delta y}{\Delta u} = f'(u) \quad \text{and} \quad \lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = \varphi'(x)$$

we arrive at the desired formula¹.

Thus, in order to obtain the derivative of the composite function $y=f[\varphi(x)]$ we should find the derivative of the function f with respect to its argument $u=\varphi(x)$ regarded as an ordinary variable and multiply the result by the derivative of the function $\varphi(x)$ with respect to the independent variable x .

Examples. (1) The function $y=(2x^2-1)^3$ can be regarded as a function of a function composed of the cubic function $y=u^3$ and the quadratic function $u=2x^2-1$. Hence, we have

$$\begin{aligned} y' &= (u^3)' u' = (u^3)' (2x^2-1)' = 3u^2 \cdot 4x = 3(2x^2-1)^2 \cdot 4x = \\ &= 48x^5 - 48x^3 + 12x \end{aligned}$$

¹ This proof is inapplicable if Δu turns into zero. It can nevertheless be shown (but we do not present the proof here) that the theorem remains valid in this case as well.

The correctness of the result is readily checked: opening the parentheses in the expression of y as a function of x we obtain $y = 8x^6 - 12x^4 + 6x^2 - 1$ whence, after the differentiation of the polynomial, we obtain $y' = 48x^5 - 48x^3 + 12x$, which coincides with the result obtained above.

(2) Consider the function $y = \left(x + \frac{1}{x}\right)^{100}$. Here the intermediate function $u = x + \frac{1}{x}$ is raised to the hundredth power. Therefore, we have

$$y' = 100u^{99}u' = 100\left(x + \frac{1}{x}\right)^{99}\left(x + \frac{1}{x}\right)' = 100\left(x + \frac{1}{x}\right)^{99}\left(1 - \frac{1}{x^2}\right)$$

This example clearly demonstrates the advantage of the application of the rule for differentiating a composite function: the direct differentiation of the given expression as a sum of different powers of x would involve Newton's binomial formula with exponent $n = 100$!

(3) Take the function $y = \sqrt{1+x^2}$. Here we can put $y = \sqrt{u}$ and $u = 1+x^2$, which yields

$$y' = \left(u^{\frac{1}{2}}\right)' u' = \frac{1}{2} u^{-\frac{1}{2}} 2x = \frac{x}{\sqrt{1+x^2}}$$

Now suppose that we are given a function of a function formed of more than two constituent functions. For definiteness, let

$$y = f(u), \quad u = \varphi(v) \quad \text{and} \quad v = \psi(x)$$

where the functions $f(u)$, $\varphi(v)$ and $\psi(x)$ possess derivatives with respect to their arguments. By the above rule, we can write

$$y' = f'(u) u'$$

where u' is the derivative of u regarded as a function of the independent variable x . The same rule also implies

$$u' = \varphi'(v) v' = \varphi'(v) \psi'(x)$$

Substituting the latter expression into the former we obtain

$$y' = f'(u) \varphi'(v) \psi'(x)$$

The general formula for an arbitrary number of intermediate arguments is derived in like manner. In all cases the resultant derivative y' is equal to the product of the derivative with respect to the first intermediate argument by the derivative with respect to the second intermediate argument and by the derivative of the third intermediate argument, etc., the terminal factor being the derivative of the last intermediate argument with respect to the independent variable x .

This can be stated briefly as follows:

The derivative of a composite function is equal to the product of the derivatives of the constituent functions.

For example, let $y = (1 + \sqrt{x^2 + 1})^2$. Representing this function as a "chain"¹ of constituent functions we write $y = u^2$, $u = 1 + \sqrt{v}$ and $v = x^2 + 1$. The differentiation now yields

$$y' = (u^2)' (1 + \sqrt{v})' (x^2 + 1)' = 2u \frac{1}{2\sqrt{v}} \cdot 2x = \frac{2(1 + \sqrt{x^2 + 1})x}{\sqrt{x^2 + 1}}$$

11. Differentiating inverse function. Let $y = f(x)$ and $x = \varphi(y)$ be a pair of mutually inverse functions (see Sec. 18). We shall show that if the derivative of one of these functions is known it is easy to determine the derivative of the other. For definiteness, let the derivative $f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}$ be known. We shall suppose that it does not turn into zero. In order to find the derivative $\varphi'(y)$ we should compute the limit $\lim_{\Delta y \rightarrow 0} \frac{\Delta x}{\Delta y}$. Since $\Delta x \rightarrow 0$ when $\Delta y \rightarrow 0$ (because the inverse function is also continuous) the identity $\frac{\Delta x}{\Delta y} = \frac{1}{\frac{\Delta y}{\Delta x}}$ implies

$$\lim_{\Delta y \rightarrow 0} \frac{\Delta x}{\Delta y} = \frac{1}{\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x}}$$

that is, $\varphi'(y) = \frac{1}{f'(x)}$. Similarly, if $\varphi'(y) \neq 0$,

then $f'(x) = \frac{1}{\varphi'(y)}$. Briefly, the derivatives

of two mutually inverse functions are the reciprocals of each other. This can be written as

$$y'_x = \frac{1}{x'_y} \quad \text{or} \quad x'_y = \frac{1}{y'_x}$$

where the subscript x or y indicates the variable with respect to which the differentiation is performed (that is, which of the variables is taken as the independent one).

The relationship between the derivatives of mutually inverse functions admits of a visual geometrical interpretation. Let the line L shown in Fig. 51 be the graph of a function $y = f(x)$. Then

$$f'(x) = \tan \alpha$$

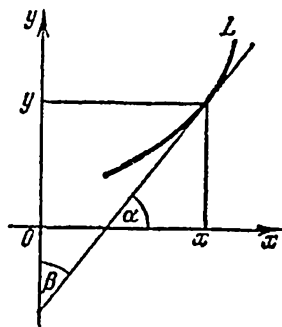


Fig. 51

¹ The rule for differentiating a composite function stated above is sometimes referred to as the "chain rule".—Tr.

The equation of the line L can also be taken in the form $x = \varphi(y)$. The derivative of this function with respect to the variable y is obviously equal to the tangent of the angle of inclination of the tangent line of the curve L to the y -axis, that is

$$\varphi'(y) = \tan \beta$$

But $\alpha + \beta = \frac{\pi}{2}$, and therefore $\tan \alpha = \cot \beta = \frac{1}{\tan \beta}$ whence

$$f'(x) = \frac{1}{\varphi'(y)}$$

Thus, this formula expresses the evident fact that *the sum of the angles formed by the tangent line with the coordinate axes is equal to $\frac{\pi}{2}$* .

As an example, take the function $y = \sqrt[3]{x}$. Here we have $x = y^3$ and $x'_y = 3y^2$. Consequently, $y' = \frac{1}{3y^2}$ (the subscript x in y' is omitted). Now let us express the derivative y' in terms of the independent variable x . To this end, we substitute into the equality obtained the expression of y as function of x to find

$$y' = \frac{1}{3\sqrt[3]{x^2}}$$

Let the reader differentiate the function $y = \sqrt[3]{x} = x^{\frac{1}{3}}$ with the aid of the general formula for differentiating a power function to check that the result is the same as above.

43. Derivatives of Basic Elementary Functions. In Sec. 39 we computed the derivative of the power function:

$$(x^n)' = nx^{n-1} \quad (1)$$

Now let us find the derivatives of all the other basic elementary functions. For brevity, the increment Δx will be denoted by h .

I. Derivatives of trigonometric functions. Let

$$y = \sin x$$

If x receives an increment $\Delta x = h$ then

$$\begin{aligned} \Delta y &= \sin(x+h) - \sin x = 2 \sin \frac{x+h-x}{2} \cos \frac{x+h+x}{2} = \\ &= 2 \sin \frac{h}{2} \cos \left(x + \frac{h}{2}\right) \end{aligned}$$

Consequently,

$$\frac{\Delta y}{h} = \frac{\sin \frac{h}{2}}{\frac{h}{2}} \cos \left(x + \frac{h}{2}\right)$$

and

$$y' = \lim_{h \rightarrow 0} \frac{\Delta y}{h} = \lim_{h \rightarrow 0} \frac{\sin \frac{h}{2}}{\frac{h}{2}} \lim_{h \rightarrow 0} \cos \left(x + \frac{h}{2} \right)$$

As was proved in Sec. 30, the first factor is equal to unity while

$$\lim_{h \rightarrow 0} \cos \left(x + \frac{h}{2} \right) = \cos (x + 0) = \cos x$$

since the cosine is a continuous function.

Hence,

$$(\sin x)' = \cos x \quad (2)$$

The differentiation formula for $y = \cos x$ can now be derived with the help of the theorem on differentiating a composite function.

For, since $\cos x = \sin \left(x + \frac{\pi}{2} \right)$, we can put $x + \frac{\pi}{2} = u$ to obtain

$$y' = (\sin u)' u' = \cos u \cdot 1 = \cos \left(x + \frac{\pi}{2} \right) = -\sin x$$

whence

$$(\cos x)' = -\sin x \quad (3)$$

We suggest that the reader derive this formula directly by computing the derivative as the limit of the ratio of the increments.

The derivatives of the functions $y = \tan x$ and $y = \cot x$ are obtained from the above formulas and from the rule for differentiating a quotient:

$$(\tan x)' = \left(\frac{\sin x}{\cos x} \right)' = \frac{(\sin x)' \cos x - (\cos x)' \sin x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}$$

and

$$\begin{aligned} (\cot x)' &= \left(\frac{\cos x}{\sin x} \right)' = \frac{(\cos x)' \sin x - (\sin x)' \cos x}{\sin^2 x} = \\ &= \frac{-\sin^2 x - \cos^2 x}{\sin^2 x} = -\frac{1}{\sin^2 x} \end{aligned}$$

Thus,

$$(\tan x)' = \frac{1}{\cos^2 x} \quad (4)$$

and

$$(\cot x)' = -\frac{1}{\sin^2 x} \quad (5)$$

Examples. (1) On differentiating $y = x \sin x$ we find

$$y' = (x)' \sin x + x (\sin x)' = \sin x + x \cos x$$

(2) Let $y = \frac{\cos x}{x}$. We obtain

$$y' = \frac{(\cos x)' x - \cos x (x)'}{x^2} = -\frac{x \sin x + \cos x}{x^2}$$

(3) Let $y = \tan 5x$. Putting $y = \tan u$ and $u = 5x$ we derive

$$y' = (\tan u)' u' = \frac{1}{\cos^2 u} \cdot 5 = \frac{5}{\cos^2 5x}$$

(4) Let $y = \sin^2 2x$. Putting $y = u^2$, $u = \sin v$ and $v = 2x$ and applying the chain rule we get

$$y' = 3u^2 \cos v \cdot 2 = 6 \sin^2 2x \cos 2x$$

II. Derivatives of inverse trigonometric functions. The derivatives of the inverse trigonometric functions are obtained from the above formulas by using the differentiation rule for inverse functions (Sec. 42).

Let $y = \arcsin x$; then $x = \sin y$ and $x'_y = \cos y$. Therefore,

$$y'_x = \frac{1}{x'_y} = \frac{1}{\cos y}$$

To express the derivative as a function of the independent variable x we substitute into this formula the expression of $\cos y$ in terms of x , i.e.

$$\cos y = \sqrt{1 - \sin^2 y} = \sqrt{1 - x^2}$$

Here we take the arithmetic root since the values of the function $y = \arcsin x$ lie in the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$, and the cosine of y is positive in this interval. The substitution gives

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} \quad (6)$$

It should be noted that although the function $y = \arcsin x$ is continuous at the points $x = \pm 1$ where it takes on the values $y = \pm \frac{\pi}{2}$ the derivative of y does not exist at these points.

Completely analogously we show that the function $y = \arccos x$ has the derivative for all the values of x in the interval $(-1, 1)$:

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}} \quad (7)$$

This result is also readily derived from the identity $\arcsin x + \arccos x = \frac{\pi}{2}$.

Now we consider the function $y = \arctan x$. Its inverse function is $x = \tan y$, and we have

$$x'_y = \frac{1}{\cos^2 y}$$

Consequently,

$$y'_x = \cos^2 y = \frac{1}{1 + \tan^2 y} = \frac{1}{1 + x^2}$$

Thus,

$$(\arctan x)' = \frac{1}{1+x^2} \quad (8)$$

and, similarly,

$$(\operatorname{arccot} x)' = -\frac{1}{1+x^2} \quad (9)$$

Examples. (1) Let us differentiate $y = x^2 \arcsin x$:

$$y' = 2x \arcsin x + \frac{x^2}{\sqrt{1-x^2}}$$

(2) For the function $y = \arctan \frac{1+x}{1-x}$ we find

$$y' = \frac{1}{1 + \left(\frac{1+x}{1-x}\right)^2} \left(\frac{1+x}{1-x}\right)' = \frac{(1-x)^2}{(1-x)^2 + (1+x)^2} \cdot \frac{2}{(1-x)^2} = \frac{1}{1+x^2}$$

The derivative obtained coincides with that of the function $\arctan x$. This is accounted for by the fact that the functions $\arctan \frac{1+x}{1-x}$ and $\arctan x$ only differ by a constant summand equal to $\frac{\pi}{4}$ if $x < 1$ and equal to $\frac{3\pi}{4}$ if $x > 1$.

III. Derivatives of logarithmic and exponential functions. Let us begin with the logarithmic function $y = \ln x$. Giving x an increment h we obtain

$$\frac{\Delta y}{h} = \frac{\ln(x+h) - \ln x}{h} = \frac{\ln \frac{x+h}{x}}{h} = \frac{\ln \left(1 + \frac{h}{x}\right)}{h}$$

This relation can be rewritten in the form

$$\frac{\Delta y}{h} = \frac{1}{x} \frac{\ln \left(1 + \frac{h}{x}\right)}{\frac{h}{x}}$$

whence it follows that

$$\lim_{h \rightarrow 0} \frac{\Delta y}{h} = \lim_{h \rightarrow 0} \left[\frac{1}{x} \frac{\ln \left(1 + \frac{h}{x}\right)}{\frac{h}{x}} \right] = \frac{1}{x} \lim_{h \rightarrow 0} \frac{\ln \left(1 + \frac{h}{x}\right)}{\frac{h}{x}}$$

According to formula (*) of Sec. 36, III, this yields

$$\lim_{h \rightarrow 0} \frac{\Delta y}{h} = \frac{1}{x}$$

Thus, the derivative of the logarithmic function $y = \ln x$ exists and is equal to $\frac{1}{x}$:

$$(\ln x)' = \frac{1}{x} \quad (10)$$

Now let $y = \log_a x$. As is known (Sec. 20, II), $\log_a x = \frac{\ln x}{\ln a}$, and therefore $(\log_a x)' = \frac{(\ln x)'}{\ln a} = \frac{1}{x \ln a}$. In particular,

$$(\log x)' = \frac{M}{x} \approx \frac{0.43429}{x} \quad \left(M = \log e = \frac{1}{\ln 10} \right)$$

To determine the derivative of the exponential function we again resort to the differentiation rule for inverse functions.

If $y = a^x$ then $\ln y = x \ln a$ and $x = \frac{\ln y}{\ln a}$ whence $x'_y = \frac{1}{y \ln a}$ and $y'_x = y \ln a = a^x \ln a$. Thus,

$$(a^x)' = a^x \ln a \quad (11)$$

In particular, for $a = e$ we obtain

$$(e^x)' = e^x \quad (12)$$

Consequently, the function $y = e^x$ remains invariable under the differentiation.

Examples. (1) For $y = x \ln x$ we obtain $y' = (x)' \ln x + x (\ln x)' = \ln x + 1$.

(2) If $y = (\ln x)^3$ then $y' = 3 (\ln x)^2 (\ln x)' = 3 \frac{\ln^2 x}{x}$.

(3) If $y = \ln \sin x$ then $y' = \frac{(\sin x)'}{\sin x} = \frac{\cos x}{\sin x} = \cot x$.

(4) Take $y = e^{-x}$. Putting $y = e^u$ and $u = -x$ we get $y' = (e^u)' u' = -e^{-x}$.

IV. Derivatives of hyperbolic and inverse hyperbolic functions. In accordance with the definitions given in Sec. 23 we obtain:

$$(\sinh x)' = \left(\frac{e^x - e^{-x}}{2} \right)' = \frac{(e^x)' - (e^{-x})'}{2} = \frac{e^x - (-e^{-x})}{2} = \cosh x$$

$$(\cosh x)' = \left(\frac{e^x + e^{-x}}{2} \right)' = \frac{(e^x)' + (e^{-x})'}{2} = \frac{e^x + (-e^{-x})}{2} = \sinh x$$

$$\begin{aligned} (\tanh x)' &= \left(\frac{\sinh x}{\cosh x} \right)' = \frac{(\sinh x)' \cosh x - \sinh x (\cosh x)'}{\cosh^2 x} = \\ &= \frac{\cosh^2 x - \sinh^2 x}{\cosh^2 x} = \frac{1}{\cosh^2 x} \end{aligned}$$

In these formulas we again see an analogy between the hyperbolic and trigonometric functions.

The derivatives of the inverse hyperbolic functions are found with the aid of the general rule for differentiating inverse func-

tions or by using the expressions of the inverse hyperbolic functions in terms of the logarithmic function (see Sec. 23). Leaving the derivation of these formulas to the reader we confine ourselves to writing down the ultimate results:

$$(\sinh^{-1} x)' = (\ln(x + \sqrt{x^2 + 1}))' = \frac{1}{\sqrt{x^2 + 1}}$$

$$(\cosh^{-1} x)' = (\ln(x \pm \sqrt{x^2 - 1}))' = \pm \frac{1}{\sqrt{x^2 - 1}}$$

(where the signs $\pm$ correspond to the two single-valued branches) and

$$(\tanh^{-1} x)' = \left(\frac{1}{2} \ln \frac{1+x}{1-x} \right)' = \frac{1}{1-x^2}$$

44. Differentiating Elementary Functions. Examples. After the general differentiation rules and the formulas for the derivatives of the basic elementary functions have been established we can readily find the derivative of any elementary function without resorting to the computation of limits.

For the sake of convenience, here we write down once again the differentiation rules and formulas (which should be learnt by heart).

Differentiation rules:

$$(u + v + \dots + w)' = u' + v' + \dots + w'$$

$$(uv)' = u'v + v'u$$

$$\left(\frac{u}{v} \right)' = \frac{u'v - v'u}{v^2}, \quad \left(\frac{1}{v} \right)' = -\frac{v'}{v^2}$$

$$\text{If } y = f(u) \text{ and } u = \varphi(x) \text{ then } y' = f'(u) \varphi'(x).$$

Differentiation formulas:

$$(1) (x^n)' = nx^{n-1}, \quad (7) (\arccos x)' = -\frac{1}{\sqrt{1-x^2}},$$

$$(2) (\sin x)' = \cos x, \quad (8) (\arctan x)' = \frac{1}{1+x^2},$$

$$(3) (\cos x)' = -\sin x, \quad (9) (\operatorname{arccot} x)' = -\frac{1}{1+x^2},$$

$$(4) (\tan x)' = \frac{1}{\cos^2 x}; \quad (10) (\ln x)' = \frac{1}{x}$$

$$(5) (\cot x)' = -\frac{1}{\sin^2 x}, \quad (11) (a^x)' = a^x \ln a$$

$$(6) (\arcsin x)' = \frac{1}{\sqrt{1-x^2}}, \quad (12) (e^x)' = e^x$$

Examples. (1) Let $y = e^{\sin(2x+1)}$. Putting $y = e^u$, $u = \sin v$ and $v = 2x+1$ we receive

$$y' = (e^u)' (\sin v)' (2x+1)' = e^u \cos v \cdot 2 = 2e^{\sin(2x+1)} \cos(2x+1)$$

After some experience in the differentiation has been gained the reader may omit the notation of intermediate arguments and introduce them mentally as a chain of functions leading from y to x .

(2) For $y = x \sin 3x$ we obtain

$$y' = \sin 3x + x (\sin 3x)' = \sin 3x + 3x \cos 3x$$

(3) If $y = \frac{e^{x^2}}{x}$ then $y' = \frac{(e^{x^2})' x - e^{x^2}}{x^2} = \frac{e^{x^2} (2x^2 - 1)}{x^2}$

45. Further Remarks on Differentiation of Functions.

1. **Logarithmic differentiation.** A function of the form

$$y = [f(x)]^{\varphi(x)} \quad (f(x) > 0)$$

where the base and the exponent are both functions of the independent variable x is referred to as a *composite exponential* (or *power-exponential*) *function*. A simple example of such a function is

$$y = x^x, \quad x > 0$$

The latter function can be differentiated as a function of a function if we write it down in the form

$$x^x = (e^{\ln x})^x = e^{x \ln x}$$

whence follows

$$(x^x)' = e^{x \ln x} (x \ln x)' = x^x (\ln x + 1)$$

But there is another way of finding this derivative. We can first take the logarithm of the function and then differentiate the resultant relation with respect to x bearing in mind that the left-hand member of the differentiated relation is a composite function of x :

$$\ln y = x \ln x$$

and

$$\frac{y'}{y} = \ln x + 1$$

if the derivative y' exists. Now we receive

$$y' = y (\ln x + 1) = x^x (\ln x + 1)$$

We have thus derived the same expression by applying another technique.

The operation of differentiating the logarithm of a function is referred to as *logarithmic differentiation*. If $f(x)$ is a given function the result of this operation is

$$[\ln f(x)]' = \frac{f'(x)}{f(x)}$$

The latter expression is called the *logarithmic derivative of the function* $f(x)$.

Logarithmic differentiation is not only applicable to finding the derivative of a composite exponential function but also to some other problems. For instance, if we have to determine the derivative of the product

$$y = 2^x \sqrt{x^2 + 4} \sin^2 x$$

it is convenient to apply logarithmic differentiation, which enables us to find the result speedily. Indeed, we have

$$\ln y = x \ln 2 + 2 \ln \sin x + \frac{1}{2} \ln (x^2 + 4)$$

whence

$$y' = y \left(\ln 2 + 2 \cot x + \frac{x}{x^2 + 4} \right)$$

Using logarithmic differentiation we can easily prove the validity of the formula

$$(x^n)' = nx^{n-1}$$

for any exponent n (by now we have proved it only for integral n 's; see Sec. 39, III). Taking the logarithm of the function $y = x^n = e^{n \ln x}$ and assuming that $x > 0$ we find

$$\ln y = n \ln x$$

On differentiating, we receive $\frac{y'}{y} = \frac{n}{x}$ whence

$$y' = \frac{n}{x} x^n = nx^{n-1}$$

If $x < 0$ then for the exponent n for which the function $y = x^n$ is defined we can write $y = (-1)^n (-x)^n$ and apply the differentiation rule for a composite function; this again leads to the required formula $y' = nx^{n-1}$.

II. Derivatives of implicit functions. Suppose that y is an implicit function of x , which means that it is specified by an equation connecting the independent variable x and the function y , the equation not being solved with respect to y (see Sec. 12). Then the derivative of this function (provided it exists) can usually be found by differentiating (with respect to x) both sides of the equation. In this differentiation it is necessary to take into account that y is a function of x (specified by this equation). We shall illustrate by examples the practical significance of this rule¹.

¹ The differentiation of implicit functions will be discussed in more detail in Sec. 118.

Let us determine the derivative of the function y specified by the equation

$$x^2 + y^2 = 1$$

(this is the equation of the circle of unit radius with centre at the origin). Differentiating with respect to x and taking into account that y is a function of x we receive

$$2x + 2yy' = 0$$

whence

$$y' = -\frac{x}{y}$$

In this example it is not difficult to find the explicit expression for y , namely $y = \sqrt{1-x^2}$ or $y = -\sqrt{1-x^2}$. The differentiation of the first function yields $y' = -\frac{x}{\sqrt{1-x^2}}$ which coincides with the result just obtained. Putting $y = -\sqrt{1-x^2}$ we similarly obtain the derivative of the branch specified by the condition $y < 0$.

Consider another example. Let the function $y = f(x)$ be specified by the equation

$$y^3 - 3axy + x^3 = 0$$

Differentiating with respect to x we get

$$3y^2y' - 3ay - 3axy' + 3x^2 = 0$$

whence

$$y' = \frac{ay - x^2}{y^2 - ax}$$

In this case it is difficult to express the derivative solely in terms of the independent variable x . It is sometimes possible to simplify the expression of the derivative of an implicit function by using the original equation connecting y and x . For instance, let us take the implicit function specified by the equation $xy^2 = e^{x+y}$. The differentiation with respect to x leads to $y^2 + 2xyy' = e^{x+y}(1+y')$ whence $y' = \frac{e^{x+y} - y^2}{2xy - e^{x+y}}$. Now, replacing e^{x+y} by xy^2 we simplify the expression of the derivative:

$$y' = \frac{y(x-1)}{x(2-y)}$$

Thus, the derivative of any implicit function specified by an equation involving elementary functions can be determined according to the known differentiation rules irrespective of whether it is possible to represent the function explicitly. In the general case the derivative of such a function is expressed in terms of the independent variable and the function itself.

46. Parametric Representation of Functions. Parametric Differentiation.

I. Parametric representation of functions and curves

Let us take two functions of one and the same variable t . We shall denote them by x and y :

$$x = \varphi(t), \quad y = \psi(t) \quad (*)$$

The specification of these functions yields a functional relationship between the variables x and y . For with each value of t (belonging to the given domain) the system (*) associates some values of x and y and thus generates a correspondence between x and y .

Definition. If a functional relationship between two variables is specified so that each variable is determined separately as a function of one and the same auxiliary variable we say that this functional relationship is represented *parametrically*, and call the auxiliary variable a *parameter*.

The determination of the direct relationship between the variables x and y (not involving the parameter t) from system (*) is referred to as the *elimination of the parameter*. This results in an equation connecting x and y which specifies one of the variables as a function of the other.

The general rule for eliminating the parameter states as follows: we take one of the given equalities, for definiteness the first, and determine from it the expression of t in terms of x , that is, find a function $t = \chi(x)$ inverse to the function φ , and substitute this expression into the other equality:

$$y = \psi[\chi(x)]$$

We thus obtain an explicit expression of y as a function of x . In concrete problems some other techniques are also used (see II, Example 1).

Interpreting the parametric relationship between x and y as specifying a curve in the xy -plane we say that this curve is represented *parametrically*, the equalities $x = \varphi(t)$, $y = \psi(t)$ being termed *parametric equations* of that curve.

A parametric specification of a function (or of a curve) is sometimes more convenient than other ways of representing the function because parametric equations expressing a complicated relationship between x and y may be considerably simpler than the original equation. Besides, in a parametric representation it is not preassigned which of the variables is the argument and which is the function.

It should be noted that if at least one of the functions in system (*) is constant it is impossible to eliminate the parameter t to obtain a function $y = f(x)$. For instance, the equations

$x=1$, $y=\sin t$ do not provide a functional relationship between x and y . Here the points with coordinates x and y corresponding to different values of t lie on the line segment $x=1$ joining the points $(1, -1)$ and $(1, 1)$.

A concrete interpretation of the parameter depends on the character of the functional relationship in question.

Functions represented parametrically are often used in mechanics in connection with the specification of the *trajectory of motion*. In this case the role of the parameter is played by time t , the functions $x(t)$ and $y(t)$ expressing the law of motion of the projections of the moving point on the coordinate axes. On eliminating the parameter t we obtain an equation for the trajectory connecting the coordinates x and y of the moving point.

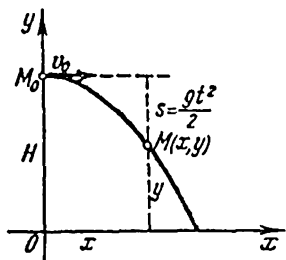


Fig. 52

Let us discuss a simple example. Suppose that a body is dropped from an airplane in a horizontal flight at a height H with velocity v_0 . We shall take the coordinate system shown in Fig. 52 and assume that at the initial time moment $t=0$ the body (together with the airplane) is at the point M_0 . Let the body be at the point $M(x, y)$ at time t . Since there are no horizontal forces acting upon the body (air resistance is neglected¹) the horizontal component of the velocity of the body is constant and equal to v_0 . Therefore, the horizontal displacement during time t is $x=v_0t$. The vertical displacement of the body corresponding to that time t is equal to $s=\frac{gt^2}{2}$ where g is the acceleration of gravity, and hence $y=H-\frac{gt^2}{2}$ (see Fig. 52). We have thus obtained the parametric equation of the trajectory:

$$y=H-\frac{gt^2}{2}, \quad x=v_0t$$

On eliminating the parameter t we arrive at the equation

$$y=H-\frac{g}{2v_0^2}x^2$$

This equation shows that the trajectory is a *parabola* with the point M_0 as *vertex*.

II. Parametric equations of some curves.

(1) Consider the circle with centre at the origin and radius a . Let t be the arc of the circle (expressed in radians) connecting

¹ The inclusion of air resistance leads to a considerable complication of the problem.

the point $(a, 0)$ and the moving point (x, y) . It is obvious that

$$x = a \cos t, \quad y = a \sin t$$

The latter equations provide a parametric representation of the circle. To eliminate t we square both equations and then add them together termwise, which results in

$$x^2 + y^2 = a^2$$

As t runs over the interval $[0, 2\pi]$ the point with coordinates x, y makes one circuit along the circle starting with the point $(a, 0)$.

Let us check that the equations

$$y = a \frac{2t}{1+t^2}, \quad x = a \frac{1-t^2}{1+t^2}$$

specify parametrically the same circle (the geometric significance of the parameter t in the latter equations is mentioned below). Squaring both equations and adding them together we again receive

$$x^2 + y^2 = a^2$$

When the parameter t ranges from 0 to ∞ the variable point traverses the upper semicircle from the point $(a, 0)$ to the point $(-a, 0)$, and if t varies from 0 to $-\infty$ it describes the lower semicircle. Hence, if t continuously varies from $-\infty$ to $+\infty$ the moving point makes one circuit over the circle traversing it counterclockwise, the initial point being $(-a, 0)$. Let the reader show that in this case the role of the parameter t is played by the tangent of half the arc of the circle joining the point $(a, 0)$ to the variable point (x, y) , the whole arc being equal to the value of the parameter in the previous parametric representation.

This example shows that a given line can be specified parametrically in various ways.

(2) Let us take the ellipse with centre at the origin and semi-axes a and b along the coordinate axes Ox and Oy . Draw the circle with the same centre and radius a (we suppose that $a > b$). Now, taking as the parameter t the arc of that circle joining the point $(a, 0)$ to the point of the circle whose abscissa coincides with that of the moving point (x, y) of the ellipse we arrive at the parametric equations of the ellipse

$$x = a \cos t, \quad y = b \sin t$$

For, on eliminating t from these equations we obtain the well-known canonical (standard) equation of the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

t being the magnitude of $\angle MCP$ measured in radians. Thus, $x = a(t - \sin t)$. Similarly, for the ordinate y of the point M we obtain

$$y = NM = PC - LC = a - a \cos t = a(1 - \cos t)$$

Consequently, the parametric equations

$$x = a(t - \sin t), \quad y = a(1 - \cos t)$$

specify the cycloid. One arc of the cycloid connecting two consecutive points of contact with Ox is traversed as t runs through the interval $[0, 2\pi]$.

The elimination of the parameter t leads to a rather complicated relationship between x and y :

$$x \pm \sqrt{y(2a - y)} = a \arccos \frac{a - y}{a}, \quad 0 \leq x \leq 2\pi a$$

III. Differentiating functions represented parametrically. Here we shall study a method for differentiating a function specified by parametric equations (*parametric differentiation*). Let y as function of x be represented parametrically by equations

$$x = \varphi(t), \quad y = \psi(t)$$

Differentiating the second of these functions with respect to x according to the differentiation rule for composite functions we derive

$$y'_x = \psi'(t) t'_x$$

where the subscript indicates the variable considered as independent with respect to which the differentiation is performed. The derivative t'_x is found with the aid of the differentiation rule for an inverse function:

$$t'_x = \frac{1}{x'_t} = \frac{1}{\varphi'(t)}$$

Finally,

$$y'_x = \frac{\psi'(t)}{\varphi'(t)} \quad (\varphi'(t) \neq 0)$$

or, in the contracted notation, $y'_x = \frac{y'_t}{x'_t}$.

As an example, let us determine the slope of the tangent line to the circle $x = a \cos t$, $y = a \sin t$. The application of the above formula yields the expression

$$y'_x = \frac{y'_t}{x'_t} = \frac{a \cos t}{-a \sin t} = -\cot t$$

The slope of the radius drawn to the moving point of the circle being apparently equal to $\tan t$, we see that the expression obtained simply expresses the well-known fact that the tangent to a circle is perpendicular to its radius.

§ 3. Some Geometrical Problems. Graphical Differentiation

47. Tangent and Normal to a Line. If we are given the equation $y=f(x)$ of a line l in Cartesian coordinates Oxy in the plane the geometrical meaning of the derivative makes it possible to solve purely analytically, that is, solely with the aid of calculations, geometrical problems connected with the *tangent* to that line. To begin with, let us form the equation of the tangent line to the curve l at its point $M_0(x_0, y_0)$ where $y_0=f(x_0)$. Since the tangent line passing through the point $M_0(x_0, y_0)$ has the slope equal to $f'(x_0)$ (Sec. 40) its equation can be written in the form

$$y - y_0 = f'(x_0)(x - x_0) \quad (*)$$

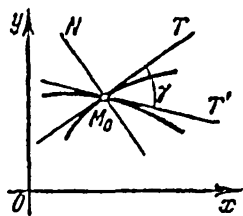


Fig. 54

For instance, the equation of the tangent line to the graph of the function $y=e^{2x}$ at its point $(0, 1)$ is written as $y-1=(2e^{2x})_{x=0}x$, that is $y=2x+1$.

Now we proceed to the definition of the normal to a curve l (see Fig. 54).

Definition. The straight line M_0N passing through the point M_0 perpendicularly to the tangent M_0T at that point is called the *normal to the curve* at its point M_0 .

The slope of the normal to the line l at the point $M_0(x_0, y_0)$ being equal to $-\frac{1}{f'(x_0)}$, the equation of the normal can be written in the form

$$y - y_0 = -\frac{1}{f'(x_0)}(x - x_0) \quad (**)$$

From analytical geometry we know the formula for computing the angle between two intersecting straight lines. Let us introduce the general definition of the angle between two intersecting curves (see Fig. 54).

Definition. The *angle γ between two curves intersecting at a point M_0* is the angle TM_0T' between the tangents drawn to the curves at the point of intersection.

If both curves have a common tangent at the point M_0 we say that they are *tangent* to each other at that point.

Below we consider some examples of the application of the derivative to solving geometrical problems.

Examples. (1) Let us take the parabola $y=ax^2$. The equation of the tangent to this line at its point $M_0(x_0, y_0)$ has the form

$$y - y_0 = 2ax_0(x - x_0)$$

But $y_0 = ax_0^2$, and therefore $y - ax_0^2 = 2ax_0(x - x_0)$, that is

$$y = 2ax_0x - ax_0^2 = 2ax_0\left(x - \frac{x_0}{2}\right)$$

It follows from this equation that the tangent meets Ox when $x = \frac{x_0}{2}$, that is, at the midpoint S of the segment ON (Fig. 55).

We have thus established a simple rule for constructing the tangent to the parabola.

Here we shall also prove an important property of the parabola: the normal M_0Q (see Fig. 55) at every point M_0 of the parabola bisects the angle formed by the focal radius FM_0 and the straight line M_0M_1 parallel to the axis of the parabola, that is $\angle QM_0F = \angle QM_0M_1$

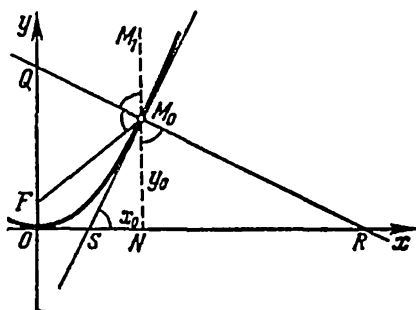


Fig. 55

(Fig. 55). It is known from analytical geometry that $OF = \frac{1}{4a}$.

The slope of the straight line FM_0 is equal to $\frac{y_0 - \frac{1}{4a}}{x_0}$. According to the above, the slope of the normal M_0Q is equal to $-\frac{1}{2ax_0}$. Now, applying the formula for the tangent of the angle between two straight lines we obtain

$$\tan \angle QM_0F = \frac{\frac{y_0 - \frac{1}{4a}}{x_0} + \frac{1}{2ax_0}}{1 - \left(\frac{y_0 - \frac{1}{4a}}{x_0}\right) \frac{1}{2ax_0}}$$

Replacing y_0 by ax_0^2 we derive, after simple transformations,

$$\tan \angle QM_0F = 2ax_0 = y'_{x=x_0}$$

Consequently, $\angle QM_0F$ is equal to the angle of inclination of the tangent line M_0S to the axis of abscissas, that is to $\angle NSM_0$. On the other hand,

$$\angle QM_0M_1 = \angle NM_0R = \angle NSM_0$$

Hence,

$$\angle QM_0F = \angle QM_0M_1$$

which is what we set out to prove.

This property of the parabola is widely used in engineering. For instance, the reflector of a projector is shaped in the form of a surface of revolution generated by the rotation of a parabola about its axis (the *parabolic reflector*) because, according to the above property (termed the *optical property*), if the light propagates in the plane of the parabola from a light source placed in its focus all the rays reflected by the parabola are parallel to its axis. Indeed, according to the well-known *law of reflection*, the angle of incidence is equal to the angle of reflection.

(2) Let us derive the equation of the tangent to an ellipse at its point $M_0(x_0, y_0)$, the ellipse being specified by its standard equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

On differentiating the implicit function y specified by this equation we obtain

$$\frac{2x}{a^2} + \frac{2y}{b^2} y' = 0$$

whence

$$y' = -\frac{b^2}{a^2} \frac{x}{y}$$

Thus, the slope of the tangent at the point $M_0(x_0, y_0)$ is equal to $-\frac{b^2}{a^2} \frac{x_0}{y_0}$.

The equation of the tangent line is now written in the form

$$y - y_0 = -\frac{b^2}{a^2} \frac{x_0}{y_0} (x - x_0)$$

which, after some simple transformation, leads to

$$\frac{y_0 y}{b^2} - \frac{y_0^2}{b^2} = -\frac{x_0 x}{a^2} + \frac{x_0^2}{a^2}$$

Since $\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1$, the desired equation of the tangent reduces to

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} = 1$$

Similarly, the equation of the tangent to the hyperbola $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$ at its point $M_0(x_0, y_0)$ is written as

$$\frac{x_0 x}{a^2} - \frac{y_0 y}{b^2} = 1$$

(3) Let us take the parametric equations of a cycloid

$$x = a(t - \sin t), \quad y = a(1 - \cos t)$$

(see Sec. 46, II) and find the equation of the tangent line at its point $M_0(x_0, y_0)$ corresponding to a value t_0 of the parameter t . The slope of the tangent at the point M_0 is equal to

$$y' = \frac{a \sin t_0}{a(1 - \cos t_0)} = \cot \frac{t_0}{2}$$

Consequently, the equation of the tangent is

$$y - y_0 = \cot \frac{t_0}{2} (x - x_0)$$

The abscissa of the point P (see Fig. 56) is $x = at_0$. For this point we obtain

$$y = a(1 - \cos t_0) + \cot \frac{t_0}{2} [at_0 - a(t_0 - \sin t_0)] = 2a$$

whence it follows that the tangent to a cycloid passes through the upper point T of the rolling circle generating the cycloid

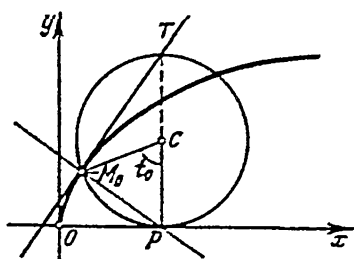


Fig. 56

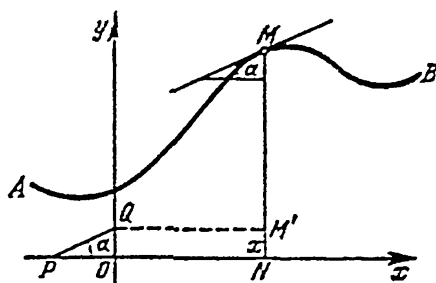


Fig. 57

while the normal passes through the lower point P (when the circle occupies the position corresponding to the given point of tangency M_0).

48. **Graphical Differentiation.** The graph of a function is used for finding the derivative mainly when the analytic expression of the function is not known and it is only specified graphically (for instance, with the aid of a self-recording apparatus).

The determination of the derivative of a function by means of its graph is referred to as *graphical differentiation*.

Suppose that the curve AB shown in Fig. 57 is the graph of a function $y = f(x)$. Let us lay off on the axis of abscissas to the left of the origin a line segment OP of unit length. Next, we draw "by eye" the tangent to the line AB at the point M corresponding to the given value x of the abscissa, and also draw from the point P the straight line parallel to that tangent. This straight line cuts the axis of ordinates at a point Q , and the length of

the line segment OQ is equal to the corresponding value of the derivative $f'(x)$. Indeed, we have

$$OQ = OP \tan \alpha = 1 \cdot \tan \alpha = f'(x)$$

Such an approximate construction of the tangent line does not provide a sufficient accuracy. The tangent can be constructed more precisely if we use a simple device, a shiny (metallic) ruler or a rectangular mirror. The reflecting surface is applied perpendicularly to the plane of the graph at the point M and then is turned about M until the reflection of the graph forms its extension without a corner point. In this position the edge of the ruler (or mirror) is in the direction of the normal at the point M , and the perpendicular to it goes along the tangent.

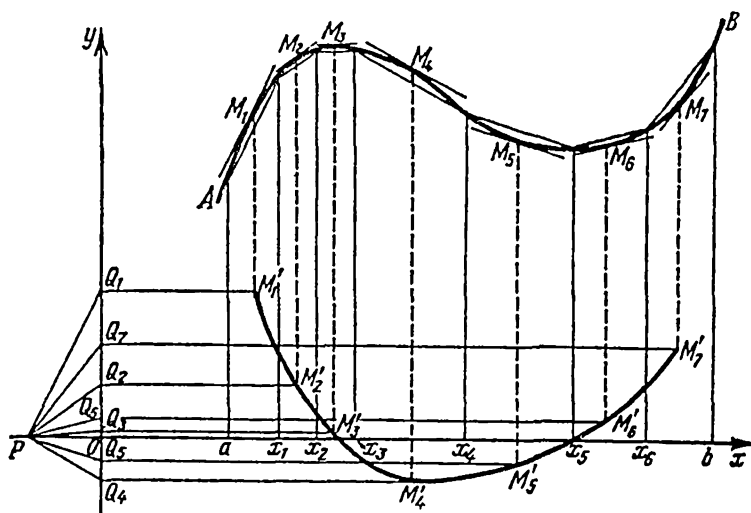


Fig. 58

After the point Q has been constructed we draw through it the straight line parallel to Ox to intersect the ordinate MN or its extension at a point M' . The ordinate of the point M' is equal to the value of the derivative $f'(x)$ for the given value of the independent variable x . To trace the graph of the derivative $y = f'(x)$ corresponding to the given graph of the function $y = f(x)$ we break up the arc AB of the curve $y = f(x)$ (Fig. 58) corresponding to the interval of variation of the independent variable x we are interested in into a number of parts by straight lines $x = x_1, x = x_2, \dots$. Then we determine graphically the values of the derivative at the midpoints of the subintervals. It is convenient to use the midpoints of the subintervals since, as a rule, the directions of the tangents at these points can be taken, with a

high accuracy, as parallel to the chords joining the end points of the arcs corresponding to the subintervals.

On constructing the points $M'_1, M'_2, \dots$ of the graph of the derivative we draw through them a continuous line which is an approximation to the graph of the derivative $y=f'(x)$. The greater the number of the intermediate points M' (i.e. the greater the number of parts the whole interval is divided into), the higher the accuracy of the construction of the graph. These parts are not necessarily taken with equal lengths; their lengths should be chosen so that the corresponding arcs are as close in their shape to line segments as possible. Therefore, those intervals where the graph of $y=f(x)$ changes its direction sharply and frequently should be broken into a greater number of parts so that every such part is sufficiently small.

We suppose that the scales along the axes of abscissas and ordinates are the same; this condition is necessary for the relation $f'(x)=\tan \alpha$ to hold.

49. Geometrical Interpretation of the Derivative in Polar Coordinates. So far we considered equations of curves in Cartesian coordinates, the independent variable and the function being, respectively, regarded as the *abscissa* and the *ordinate* of a point in the plane. As was shown, in this interpretation the derivative y' of the function $y=f(x)$ is equal to the slope of the tangent to the graph of the function.

Now let us interpret the independent variable and the function as the *polar angle* and the *polar radius* of a point in the coordinate plane. The curve specified by an equation of the form

$$r=f(\varphi)$$

connecting the independent variable φ and the function r is called the *graph of the function* $r=f(\varphi)$ in polar coordinates.

The geometrical significance of the derivative $r'=f'(\varphi)$ is elucidated by the following theorem:

Theorem. The value of the derivative $r'=f'(\varphi)$ is equal to the length of the polar radius multiplied by the cotangent of the angle θ between the polar radius and the tangent drawn to the graph of the function $r=f(\varphi)$ at the point with polar angle φ :

$$r'=r \cot \theta$$

Proof. Let us take a value φ of the argument and give it an increment $\Delta\varphi$. The values φ and $\varphi+\Delta\varphi$ determine the corresponding points M and M' of the graph of the function $r=f(\varphi)$ (Fig. 59). The length of the polar radius PM is equal to the

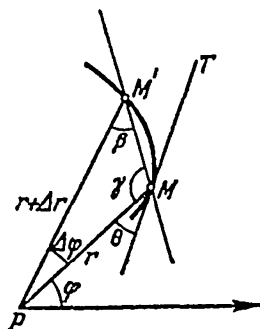


Fig. 59

given value of the function and that of the polar radius PM' to the value $r + \Delta r = f(\varphi + \Delta\varphi)$ of the function. To find Δr we apply to the triangle PMM' the law of sines:

$$\frac{r + \Delta r}{r} = \frac{\sin \gamma}{\sin \beta} = \frac{\sin [\pi - (\beta + \Delta\varphi)]}{\sin \beta} = \frac{\sin (\beta + \Delta\varphi)}{\sin \beta}$$

This yields

$$1 + \frac{\Delta r}{r} = \frac{\sin \beta \cos \Delta\varphi + \cos \beta \sin \Delta\varphi}{\sin \beta} = \cos \Delta\varphi + \cot \beta \sin \Delta\varphi$$

whence

$$\frac{\Delta r}{r} = \cos \Delta\varphi - 1 + \cot \beta \sin \Delta\varphi = -2 \sin^2 \frac{\Delta\varphi}{2} + \cot \beta \sin \Delta\varphi$$

The division of both sides of this equality by $\Delta\varphi$ leads to

$$\frac{1}{r} \frac{\Delta r}{\Delta\varphi} = -2 \frac{\sin^2 \frac{\Delta\varphi}{2}}{\Delta\varphi} + \cot \beta \frac{\sin \Delta\varphi}{\Delta\varphi}$$

If $\Delta\varphi \rightarrow 0$ the first term on the right-hand side of the latter relation tends to zero while the second term to $\lim_{\Delta\varphi \rightarrow 0} \cot \beta$, the left-hand side approaching $\frac{1}{r} r'$. Consequently,

$$\frac{1}{r} r' = \lim_{\Delta\varphi \rightarrow 0} \cot \beta$$

When $\Delta\varphi$ tends to zero the point M' approaches the point M , and, in the limit, the secant MM' turns to coincide with the tangent MT . Consequently, $\beta \rightarrow \theta$ as $\Delta\varphi \rightarrow 0$, and since $\cot \beta$ is a continuous function, this yields the desired relation

$$r' = r \cot \theta$$

This relation makes it possible to use differentiation for solving problems on tangents and normals to curves specified by equations in polar coordinates.

It should be noted that if, as the polar angle φ increases, the radius r increases, θ is understood as the acute angle (see Fig. 59), and if r decreases it is understood

as the obtuse angle between the polar radius and the tangent.

Example. Let us prove that the *logarithmic spiral* $r = ae^{m\varphi}$ intersects every straight line passing through the pole at a constant angle (Fig. 60). Indeed, we have $r' = ame^{m\varphi} = mr$, and therefore

$$\cot \theta = \frac{r'}{r} = m$$

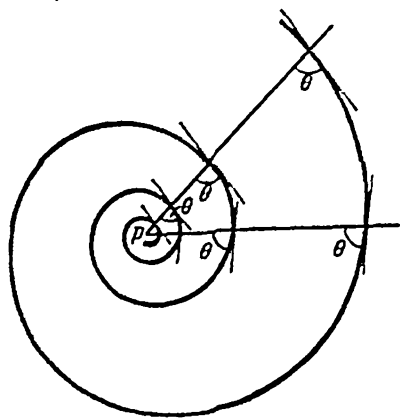


Fig. 60

§ 4. Differential

50. Differential and Its Geometrical Meaning. The notion of the derivative is closely related to another fundamental notion of mathematical analysis, namely to that of the differential of a function.

Let, for the values of x in question, $y=f(x)$ be a continuous function possessing the derivative

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x)$$

This implies

$$\frac{\Delta y}{\Delta x} = f'(x) + \epsilon$$

where ϵ is an infinitesimal as $\Delta x \rightarrow 0$. It follows that

$$\Delta y = f'(x) \Delta x + \epsilon \Delta x, \text{ i.e. } \Delta y = f'(x) \Delta x + o(\Delta x)$$

where $o(\Delta x)$ is an infinitesimal of higher order than Δx as $\Delta x \rightarrow 0$ (see Sec. 36, I).

Hence, an infinitesimal increment Δy of a function $y=f(x)$ possessing the derivative $f'(x)$ is representable as the sum of two terms which are (1) the quantity $f'(x)\Delta x$ proportional to the infinitesimal increment Δx of the independent variable, and (2) an infinitesimal of higher order than Δx .

As was shown, if $y=f(x)$ is a linear function of the form $y=ax+b$ its increment is exactly proportional to the increment Δx of the argument:

$$\Delta y = a \Delta x$$

where the proportionality factor a is constant. Now we see that if the function $f(x)$ is nonlinear and possesses the derivative $f'(x)$ its increment Δy is not exactly proportional to Δx , but nevertheless there exists a proportionality factor equal to $f'(x)$ such that the quantity $f'(x)\Delta x$ differs from the increment Δy by an infinitesimal of higher order than Δx ; in the general case this proportionality factor $f'(x)$ varies together with x .

The converse is also true: if for a given value x of the argument the increment $\Delta y = f(x+\Delta x) - f(x)$ is representable in the form

$$\Delta y = a \Delta x + o(\Delta x)$$

the function $y=f(x)$ possesses the derivative $f'(x)$, and $a=f'(x)$.

For we have

$$\frac{\Delta y}{\Delta x} = a + \frac{o(\Delta x)}{\Delta x}$$

and, since $\frac{o(\Delta x)}{\Delta x} \rightarrow 0$ as $\Delta x \rightarrow 0$, the limit of $\frac{\Delta y}{\Delta x}$ exists and is equal to a , which is what we had to prove.

Now let us introduce the following

Definition. The magnitude proportional to the infinitesimal increment Δx of the argument and differing from the corresponding increment of the function by an infinitesimal of higher order than Δx is called the *differential of that function*.

The differential of a function $y=f(x)$ is denoted by dy or $df(x)$.

We see that the differential of a function $y=f(x)$ of one independent variable exists if and only if the derivative $f'(x)$ exists, and the differential and the derivative are connected by the relationship

$$dy = f'(x) \Delta x$$

The increment Δx of the independent variable x is called its *differential* and is denoted by dx :

$$\Delta x = dx$$

This is coherent with the general definition of a differential since for the function $y=x$ we have

$$dy = dx = (x') \Delta x = \Delta x$$

that is $dx = \Delta x$.

Thus,

the differential of a function is equal to its derivative multiplied by the differential of the independent variable:

$$dy = f'(x) dx$$

Knowing the derivative we can readily find the differential and vice versa; that is why the operations of finding the derivative and the differential of a function both are termed the *differentiation*.

It follows that the derivative can be written as¹

$$f'(x) = \frac{dy}{dx}$$

This representation of the derivative as the ratio of the differentials is extremely important for mathematical analysis, which will be clearly seen later on.

¹ Our investigation shows that

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = \frac{dy}{dx}$$

It would be wrong to interpret this limiting relation in the sense that Δy tends to dy and Δx to dx , as $\Delta x \rightarrow 0$. The correct meaning is that the ratio of the increments $\frac{\Delta y}{\Delta x}$ tends to the ratio of the differentials $\frac{dy}{dx}$ as $\Delta x \rightarrow 0$. (Note that the definition of the differential does not involve the notion of the derivative.)

In the general case, for an arbitrary x , the differential $dy = f'(x) dx$ is a function of *two* arguments x and dx (independent of each other).

Let us discuss the geometrical significance of the differential of a function $y = f(x)$ (see Fig. 61). Since $f'(x) = \tan \alpha$ the differential $dy = f'(x) dx$ is equal to the length of the line segment RT , that is,

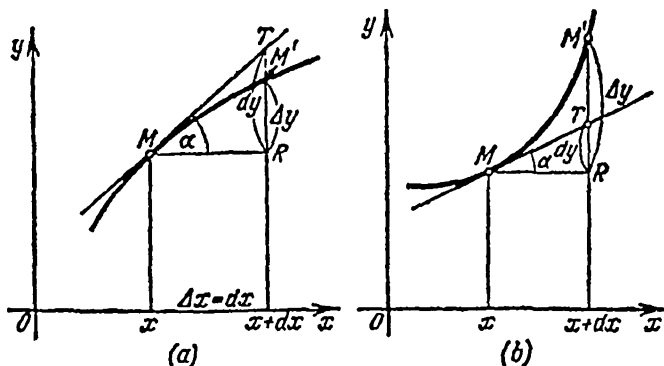


Fig. 61

the differential dy of a function $y = f(x)$ at a point x is equal to the increment of the ordinate of the tangent line drawn to the graph of the function at its corresponding point.

The increment of the function $\Delta f(x)$ is equal to the increment of the ordinate of the graph of the function (i.e. to the line segment RM' in Fig. 61), and therefore the difference between the increment of the function and its differential is equal to the length of the line segment $M'T$ lying between the tangent and the graph. *This line segment is an infinitesimal of higher order than the segment MR , as $dx \rightarrow 0$.*

For a concrete finite value of the increment Δx the differential of a function may be greater or less than its increment (see Fig. 61a and b).

51. Properties of the Differential.

I. **Differentials of basic elementary functions.** Since the differential of a function is obtained as the product of the derivative by the differential of the independent variable, we can readily write down the table of the differentials of all the basic elementary functions because their derivatives are known. For instance, $d(x^n) = nx^{n-1} dx$, $d(a^x) = a^x \ln a dx$, etc.

II. **Differentials of the results of arithmetical operations on functions.** In accordance with the rules for finding derivatives (see Sec. 41; its notation is retained here) we readily derive the following rules for computing differentials:

(a) On multiplying both sides of the relation

$$(u + v + \dots + w)' = u' + v' + \dots + w'$$

by dx we obtain

$$d(u + v + \dots + w) = du + dv + \dots + dw$$

(b) Since

$$(uv)' = u'v + uv'$$

the multiplication of both members by dx leads to

$$d(uv) = v du + u dv$$

and, in particular,

$$d(Cu) = C du$$

(c) Taking the formula

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}$$

and multiplying its both members by dx we receive

$$d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$$

III. Differential of a composite function. Invariance of the form of the differential. The extremely important invariance property of the differential follows from the differentiation rule for a function of a function.

Let $y = f(u)$ and $u = \varphi(x)$ be two continuous functions of their arguments possessing the derivatives $f'(u)$ and $\varphi'(x)$ with respect to these arguments. If we put $F(x) = f[\varphi(x)]$ then, as was shown in Sec. 42, I,

$$y' = F'(x) = f'(u) \varphi'(x) \quad (*)$$

On multiplying both sides of this relation by dx we get

$$dy = f'(u) \varphi'(x) dx$$

Since $\varphi'(x) dx = du$, this implies

$$dy = f'(u) du$$

Thus, the differential dy has the same form as if the magnitude u were an independent variable; this can be stated as follows:

the differential of a function $y = f(u)$ retains the same expression irrespective of whether its argument u is an independent variable or a function of another variable.

This property is referred to as the *invariance of the form of the differential*. It makes it possible to write down the differential in one and the same form irrespective of the nature of the argument of the function.

The equality $dy = f'(u) du$ implies

$$f'(u) = \frac{dy}{du}$$

and hence in all the cases

the rate of change of a function relative to its argument is equal to the ratio of the differential of the function to the differential of its argument.

Relation (*) can now be written as

$$\frac{dy}{dx} = \frac{dy}{du} \frac{du}{dx} \quad (**)$$

The right-hand member of equality (**) is obtained from the left-hand member by the simultaneous multiplication and division of the former by du (if, of course, $du \neq 0$). Hence, the arithmetical operations on differentials can be performed as if they were ordinary numbers. Here lies the reason for the convenience of the representation of the derivative as the ratio of the differentials. For instance, using this representation of the derivative we can readily write down the differentiation rule for the inverse of a function (cf. Sec. 42, II):

$$y'_x = \frac{dy}{dx} = \frac{1}{\frac{dx}{dy}} = \frac{1}{x'_y}$$

IV. The differential as the principal part of the increment. The definition of the differential of a function $y = f(x)$ implies the relation

$$\Delta y - dy = \Delta y - f'(x) dx = o(dx) \quad \text{as } dx \rightarrow 0$$

Let the derivative $f'(x)$ be nonzero at a point x : $f'(x) \neq 0$. Then the infinitesimals dx and dy are of the same order, and the relation $\Delta y - dy = o(dx)$ shows that

$$\Delta y - dy = o(dy)$$

which means that Δy and dy are *equivalent infinitesimals* (see Sec. 36, II):

$$\Delta y \sim dy \quad \text{for } dx \rightarrow 0$$

In what follows we shall often approximately replace the increment Δy of a function by its differential dy ; in this connection let us discuss the general question on the accuracy of the approximation to a magnitude when it is replaced by another magnitude tending to the same limit. Let $u(x)$ and $v(x)$ be two functions which tend to one and the same limit A as $x \rightarrow x_0$: $\lim u = \lim v = A$. Then the limit of their difference is equal to zero: $\lim(u - v) = 0$. Consequently, the difference $u - v$ is itself an

infinitesimal as $x \rightarrow x_0$:

$$u(x) - v(x) = \alpha(x)$$

where $\alpha(x) \rightarrow 0$ as $x \rightarrow x_0$.

If now the value of the function $u(x)$ assumed in a neighbourhood of the point x_0 is replaced by the corresponding value of the function $v(x)$ the absolute error $\alpha = |u - v|$ is an infinitesimal, and if $A \neq 0$ the relative error which characterizes the accuracy of the approximation is also an infinitesimal:

$$\delta = \frac{\alpha}{|v|} = \left| \frac{u-v}{v} \right| = \left| \frac{u}{v} - 1 \right| \rightarrow \left| \frac{A}{A} - 1 \right| = 0$$

But if $A = 0$, that is if $u(x)$ and $v(x)$ are both infinitesimals, the relative error appearing as $v(x)$ is substituted for $u(x)$ is not necessarily an infinitesimal. The results of Sec. 36 imply that in this case *for the relative error to be an infinitesimal (which means that $u - v$ is equal to $o(v)$ or to $o(u)$) it is necessary and sufficient that $u(x)$ and $v(x)$ be equivalent infinitesimals.*

This proposition (allowing us to replace one infinitesimal by another equivalent infinitesimal with an infinitesimal relative error) is often stated as follows: *each of two equivalent infinitesimals is the principal part of the other.*

As was shown, $\Delta y \sim dy$ as $dx \rightarrow 0$; therefore, using the above terminology, we can say that

the differential of a function (which is proportional to the differential (increment) of the independent variable) is the principal part of the increment of the function.

It should also be noted that if $f'(x) = 0$ at a point x the differential is equal to zero: $dy = 0$. In this case it is not compared with any infinitesimal (including the increment Δy of the function).

52. Differentiable Functions.

I. Definition. A function $y = f(x)$ is said to be *differentiable* at a point x if it possesses the differential at that point.

As was shown (see Sec. 50), the existence of the differential at the point x is equivalent to the existence of the derivative $f'(x)$ (and vice versa). Therefore, the existence of the derivative can be taken as the condition equivalent to the differentiability of the function; from the geometrical point of view, this condition is equivalent to the existence of the tangent to the curve $y = f(x)$ not perpendicular to Ox .

Now suppose that a function $f(x)$ is *differentiable* for a given value x of the argument, that is

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta y}{\Delta x} = f'(x)$$

Then it must necessarily be continuous at the point x .¹ For we

¹ This implication was already mentioned in Sec. 39.

have $\Delta y = f'(x) \Delta x + o(\Delta x)$ (see Sec. 50), and hence the increment Δy tends to zero as $\Delta x \rightarrow 0$, which means that the function $y = f(x)$ is continuous at the point x . Thus, a function has no derivative at its points of discontinuity.

Consequently, if a function $f(x)$ is differentiable it is necessarily continuous. In the general case the converse statement is not true since there are examples of continuous functions which do not possess derivatives at some points.

For instance, the function $y = |x|$ is continuous throughout Ox but has no derivative at the point $x = 0$. Indeed, we have

$$\Delta y = |x + \Delta x| - |x|$$

and the substitution of $x = 0$ yields

$$\Delta y = |\Delta x|$$

whence

$$\frac{\Delta y}{\Delta x} = \frac{|\Delta x|}{\Delta x}$$

We see that

$$\frac{\Delta y}{\Delta x} = \frac{\Delta x}{\Delta x} = 1$$

for $\Delta x > 0$ and

$$\frac{\Delta y}{\Delta x} = \frac{-\Delta x}{\Delta x} = -1$$

for $\Delta x < 0$. Therefore, the ratio $\frac{\Delta y}{\Delta x}$ has no limit as Δx tends to zero arbitrarily, which is equivalent to the nonexistence of the derivative at the point $x = 0$. This fact is also clear from the geometrical point of view: the graph of the function $y = |x|$ is a broken line (Fig. 62) with a corner point at the origin $(0, 0)$, and there is of course no tangent line to the graph at that point.

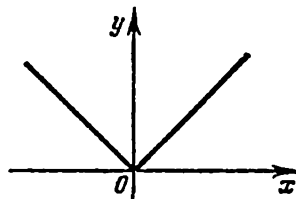


Fig. 62

It can also happen that the graph of a continuous function has a tangent at a given point but the derivative of the function does not exist at that point; this is the case when the tangent is perpendicular to the axis of abscissas. For instance, consider the function $y = \sqrt[3]{x}$; we have

$$y' = \frac{1}{3\sqrt[3]{x^2}} \quad \text{if } x \neq 0$$

and, as $x \rightarrow 0$, the derivative approaches infinity. It can be in fact shown that this function is nondifferentiable at the point

$x=0$.¹ At the point $(0, 0)$ the curve $y=\sqrt[3]{x}$ (which is the cubical parabola) touches Oy (Fig. 63). For generality, in such cases we say that the function has an *infinite derivative* although, as in the above example, it is nondifferentiable at the point in question.

The differentiability of functions and the connection between the differentiability and continuity are thoroughly investigated in

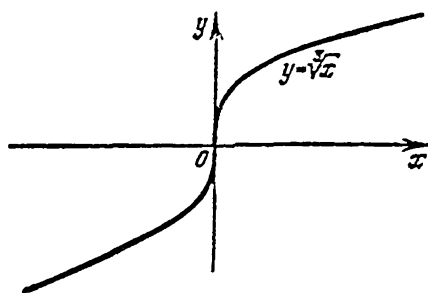


Fig. 63

mathematical analysis. The distinction between the properties of continuity and differentiability was studied by N.I. Lobachevsky². There are curious examples of continuous functions nondifferentiable at every point; the graphs of such functions are curves having no tangent at each point. They cannot of course be depicted in an ordinary manner, and it is even

difficult to imagine the structure of such a line. Functions and lines of that kind are rarely encountered in practical applications of mathematical analysis, and we shall not treat them here.

It should be noted that *all the elementary functions are differentiable everywhere in the intervals where they are defined except possibly at separate points.*

II. One-sided derivatives. Let us once again consider the function $y=|x|$ whose graph is shown in Fig. 62. As was seen, there is no finite limit of the ratio $\frac{\Delta y}{\Delta x}$ at the point $x=0$ as Δx tends to zero in an arbitrary way. At the same time, the *left-hand* and the *right-hand* limits of this ratio (see Sec. 33) exist:

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x < 0}} \frac{\Delta y}{\Delta x} = -1, \quad \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta x > 0}} \frac{\Delta y}{\Delta x} = 1$$

These limits are called, respectively, the *left-hand derivative* and the *right-hand derivative* at the point $x=0$.

¹ The derivative at the point $x=0$ cannot be investigated with the aid of the formula $y' = \frac{1}{3\sqrt[3]{x^2}}$ since it only applies to $x \neq 0$. For $x=0$ the increment

$\Delta y = \sqrt[3]{\Delta x}$ is of lower order than Δx , and therefore the ratio $\frac{\Delta y}{\Delta x}$ has no finite limit as $\Delta x \rightarrow 0$ (it tends to infinity).

² N.I. Lobachevsky (1792-1856), the great Russian mathematician, was one of the creators of non-Euclidean geometry. He contributed many outstanding results to algebra, geometry and mathematical analysis which influenced greatly further development of mathematics.

Generally, if there exist the left-hand and the right-hand limits of the ratio $\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$ they are called the *one-sided derivatives* at the point x_0 ; accordingly, for $x \rightarrow x_0$, $x < x_0$, we have the *left-hand derivative* and for $x \rightarrow x_0$, $x > x_0$, the *right-hand derivative* (we shall denote them, respectively, by $f'_l(x_0)$ and $f'_r(x_0)$). It is clear that if the left-hand and the right-hand derivatives at a given point coincide then the function is differentiable at that point and vice versa.

If a function is defined in a closed interval then, when speaking about its derivatives at the end points of the interval, we always mean the one-sided derivatives: the right-hand derivative at the left end point and the left-hand derivative at the right end point. This interpretation of differentiability at the end points of the interval of definition is analogous to that of continuity (see Sec. 32).

53. Applying Differentials to Approximate Calculations. The application of the differential to approximate calculations is based on the replacement of the increment $\Delta y = f(x_0 + \Delta x) - f(x_0)$ of a given function $y = f(x)$, which may depend on $\Delta x = dx$ in a complicated manner, by the simpler expression $f'(x_0)dx$ (the differential) found by differentiation.

Thus, for small values of dx we write

$$\Delta y = f(x_0 + dx) - f(x_0) \approx f'(x_0)dx = dy \quad (*)$$

with a small relative error (see Sec. 51, IV).

When putting $\Delta y = dy$ we replace the given function $y = f(x)$ by the *linear function*

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Geometrically, this is equivalent to replacing the graph of the function $y = f(x)$ by its tangent line at the point $(x_0, f(x_0))$. In a sufficiently small neighbourhood of the point x_0 this replacement leads to small errors. A demerit of formula (*) is that, although we know that the relative error tends to zero as $dx \rightarrow 0$, it does not provide any estimation of the error for a given numerical value of dx .¹

The approximate equality (*) can be immediately used to solve the following problem:

Given the values of $f(x_0)$, $f'(x_0)$ and dx , it is required to compute an approximation to the value $f(x_0 + dx)$ of the function.

Relation (*) directly gives us the desired formula:

$$f(x_0 + dx) \approx f(x_0) + f'(x_0)dx$$

Let us consider some illustrative examples (for brevity, we shall write x in place of x_0 and denote dx by h).

¹ Such an estimation will be derived in Chapter X.

(1) Consider the function $y = \sqrt{x}$. Its differential is $dy = \frac{1}{2\sqrt{x}} dx$, and hence

$$\sqrt{x+h} \approx \sqrt{x} + \frac{h}{2\sqrt{x}}$$

In particular, for $x=1$ we obtain

$$\sqrt{1+h} \approx 1 + \frac{h}{2}$$

In the general case, for $x=a^2$ ($a > 0$) we have

$$\sqrt{a^2+h} \approx a + \frac{h}{2a}$$

These approximate formulas are extremely simple and make it possible to compute square roots with a sufficient accuracy when $|h|$ is small compared with a^2 .

For instance, the application of these results yields $\sqrt{1.21} = \sqrt{1+0.21} \approx 1 + \frac{0.21}{2} = 1.105$. The exact value of the root is equal to 1.1; hence, the relative error is about 0.5%.

To compute the root $\sqrt{408}$ we represent it in the form $\sqrt{408} = \sqrt{20^2+8}$ and thus obtain

$$\sqrt{408} \approx 20 + \frac{8}{2 \cdot 20} = 20.2$$

The accuracy of this result is rather high since the tabular value of the root accurate to 10^{-4} is equal to 20.1990.

Now let us take $\sqrt{390}$. Here it is convenient to put $h=-10$; then

$$\sqrt{390} = \sqrt{20^2-10} \approx 20 - \frac{10}{2 \cdot 20} = 19.75$$

(the tabular value is 19.7484).

If $y = \sqrt[n]{x}$ then $dy = \frac{1}{n} \frac{\sqrt[n]{x}}{x} dx$ and $\sqrt[n]{x+h} \approx \sqrt[n]{x} + \frac{1}{n} \frac{\sqrt[n]{x}}{x} h$.

For $x=1$ this yields the approximate formula $\sqrt[n]{1+h} \approx 1 + \frac{h}{n}$. A more general formula is obtained for $x=a^n$ ($a > 0$):

$$\sqrt[n]{a^n+h} \approx a + \frac{h}{na^{n-1}}$$

Let the reader compute several roots with the aid of this formula and estimate the accuracy achieved by finding more accurate values using the table of logarithms.

(2) Let us take the function $y = \sin x$. Its differential is $dy = \cos x dx$, and therefore

$$\sin(x+h) \approx \sin x + h \cos x$$

In particular, for $x=0$ we derive the formula $\sin h \approx h$ which also follows from the equivalence of the infinitesimals $\sin h$ and h as $h \rightarrow 0$.

For example, we have $\sin \frac{\pi}{180} \approx \frac{\pi}{180} \approx 0.01745$, that is, approximately, $\sin 1^\circ = 0.01745$. This approximation is correct to the fifth decimal digit, i.e. the error does not exceed 10^{-5} .

Let us also compute $\sin 31^\circ$:

$$\sin 31^\circ \approx \sin 30^\circ + \frac{\pi}{180} \cos 30^\circ \approx 0.5 + \frac{\sqrt{3}}{2} \cdot 0.01745 \approx 0.5151$$

The tabular value of $\sin 31^\circ$ correct within 10^{-4} is equal to 0.5150. Let the reader account for the fact that we have received a major approximation of $\sin 31^\circ$.

(3) Consider the function $y = \ln x$. Here we have $dy = \frac{1}{x} dx$ and

$$\ln(x+h) \approx \ln x + \frac{h}{x}$$

In particular, for $x=1$ this yields the formula $\ln(1+h) \approx h$ which is also implied by the equivalence of the infinitesimals $\ln(1+h)$ and h as $h \rightarrow 0$ (see Sec. 36, III).

Take the known value $\ln 781 \approx 6.66058$. To compute $\ln 782$ we apply the above formulas

$$\ln 782 \approx 6.66058 + \frac{1}{781} \approx 6.66186$$

The tabular value of $\ln 782$ correct within 10^{-5} is equal to 6.66185. We see that the error of our approximation is small. Let the reader explain why in this case we have also obtained a major approximation.

Another important problem to which differentials are applied is to compute the limiting absolute error e_y of an approximation to the value of a function y for a given limiting absolute error e_x in the determination of the value of the argument x . As a rule, in practical problems the value of the argument is found by a *measurement*, and its limiting absolute error is known (see Sec. 6).

Suppose that it is required to compute the value of a given function $y=f(x)$ for a given value x of the argument whose exact value is unknown, and only an approximation to this value with limiting absolute error e_x is given:

$$x = x_0 + dx, \quad |dx| < e_x$$

Taking the value $f(x_0)$ instead of the true value $f(x)$ we obtain an approximation whose error should be estimated. From relation (*) it follows that

$$|f(x) - f(x_0)| \approx |f'(x_0)| |dx| < |f'(x_0)| e_x$$

This shows that the absolute error resulting from the replacement of $f(x)$ by the approximate value $f(x_0)$ is less than $|f'(x_0)| e_x$; the latter quantity can thus be taken as the limiting absolute error e_y of the determination of the function

$$e_y = |f'(x_0)| e_x$$

The limiting absolute error of the function is equal to the product of the absolute value of its derivative at the point corresponding to the approximate value of the argument by the limiting absolute error of the argument.

The limiting relative error δ_y of the function is given by the formula

$$\delta_y = \frac{e_y}{|f(x_0)|} = \left| \frac{f'(x_0)}{f(x_0)} \right| e_x$$

It is equal to the product of the absolute value of the logarithmic derivative by the limiting absolute error of the argument.

These formulas make it possible to estimate the admissible value of the error e_x in the determination of the argument guaranteeing the given admissible error of the value of the function:

$$e_x = \frac{e_y}{|f'(x_0)|} = \left| \frac{f(x_0)}{f'(x_0)} \right| \delta_y$$

Of course, to compute e_x we must know an approximate value x_0 of the argument which is required for finding the corresponding approximations $f(x_0)$ and $f'(x_0)$ to the function and its derivative (see Example 2).

Let us consider two illustrative examples.

(1) Given the hypotenuse c of a right triangle and an approximate value α of its acute angle with limiting absolute error e_α , it is required to determine the relative errors appearing in the computation of the legs of the triangle.

We have

$$a = c \sin \alpha, \quad b = c \cos \alpha$$

whence

$$\delta_a = \cot \alpha \cdot e_\alpha, \quad \delta_b = \tan \alpha \cdot e_\alpha$$

If $\alpha > \frac{\pi}{4}$ then $a > b$ and $\cot \alpha < \tan \alpha$; therefore the greater leg is found with a smaller relative error.

(2) The volume of a cube should be found with relative error $\delta = 0.01 = 1\%$. The edge of the cube is measured with a ruler with scale spacing of 1 cm. The approximate result of the measurement is 20 cm. Does the corresponding approximate value $v = 8000 \text{ cm}^3$ of the volume guarantee the required accuracy?

Denote the edge of the cube by x ; then $v = x^3$ and $v' = 3x^2$. By the formula for e_x we find

$$e_x = \frac{8000}{3 \cdot 400} 0.01 \approx 0.07 \text{ cm}$$

It is seen that the error in the measurement of the edge which is of 0.5 cm considerably exceeds 0.07 cm, and hence the accuracy is insufficient. It is clear that to achieve the desired accuracy we should make a repeated *measurement* with a ruler having the scale spacing of 1 mm (which provides the value $e_x = 0.05 \text{ cm}$ of the absolute limiting error in the determination of x) and compute the new value of the volume.

§ 5. Derivatives and Differentials of Higher Orders

54. Derivatives of Higher Orders.

1. Suppose that a function $y = f(x)$ defined in an interval has the derivative $f'(x)$ with respect to the independent variable x . If the function $f'(x)$ is, in its turn, differentiable its derivative is called the *second derivative* (or the *derivative of the second order*) of the original function $f(x)$ and is denoted $f''(x)$:

$$f''(x) = [f'(x)]' = \lim_{\Delta x \rightarrow 0} \frac{f'(x + \Delta x) - f'(x)}{\Delta x}$$

Similarly, the *third derivative* or the *derivative of the third order* of the function $y = f(x)$ (denoted by $f'''(x)$) is the derivative (provided it exists) of the second derivative. Now we state the general definition:

Definition. The *derivative of the n th order* $f^{(n)}(x)$ is the derivative of the derivative of the $(n-1)$ th order:

$$f^{(n)}(x) = [f^{(n-1)}(x)]' = \lim_{\Delta x \rightarrow 0} \frac{f^{(n-1)}(x + \Delta x) - f^{(n-1)}(x)}{\Delta x}$$

To unify the terminology we call the derivative $f'(x)$ the *derivative of the first order* or the *first derivative*.

Derivatives of the second order and, generally, of higher orders play an essential role in mathematics and its applications to mechanics, physics and other sciences; they are used for a more complete investigation of functions which cannot be achieved when we solely consider the first derivative.

The value $f''(x)$ of the second derivative at a point x characterizes the rate of change of $f'(x)$ at that point, that is the rate of change of the rate of change of $f(x)$. By analogy with mechanics, we can say that $f''(x)$ is the *acceleration of the change of the function $f(x)$ at the given point x* .

The geometrical meaning of the second derivative will be discussed in Chapter IV (Sec. 62).

If a point is in rectilinear motion and a functional relationship $s = F(t)$ describes the law of motion the first derivative $F'(t)$ is the velocity of motion of the point (see Sec. 38, I). Accordingly, the second derivative $F''(t)$ is the *acceleration of the rectilinear motion of the point*.

Traditionally, due to Newton, in mechanics derivatives with respect to time t are marked by dots above the symbols of functions instead of primes. In this notation the formulas for the velocity v and the acceleration w of a rectilinear motion of a point are written in the form

$$v = \dot{s} \text{ and } w = \dot{v} = \ddot{s} \text{ (where } s = F(t) \text{ is the law of motion)}$$

Since the rules and formulas for computing derivative of any elementary function are known, we can find derivatives of arbitrary orders of a given elementary function by differentiating it, in succession, the required number of times. In some cases it is possible to derive the general formula for the derivative of the k th order of a given function for an arbitrary k . For brevity, such a derivative is also denoted by $y^{(k)}$.

Examples. (1) If $y = x^n$ then $y' = nx^{n-1}$, $y'' = n(n-1)x^{n-2}$, ..., $y^{(k)} = n(n-1)(n-2)\dots(n-k+1)x^{n-k}$.

If n is a positive integer, we have

$$y^{(n)} = n! \text{ and } y^{(n+1)} = y^{(n+2)} = \dots = 0$$

(2) Taking $y = \sin x$ we find $y' = \cos x = \sin\left(x + \frac{\pi}{2}\right)$, $y'' = \sin(x + \pi)$, ..., $y^{(k)} = \sin\left(x + \frac{k\pi}{2}\right)$.

(3) For $y = e^x$ we obtain $y^{(k)} = e^x$.

(4) If $y = a^x$ then $y' = a^x \ln a$, $y'' = a^x (\ln a)^2$, ..., $y^{(k)} = a^x (\ln a)^k$.

(5) Differentiating the function $y = \ln(1+x)$ we receive

$$y' = \frac{1}{1+x}, \quad y'' = -\frac{1}{(1+x)^2}, \quad \dots$$

$$\dots, \quad y^{(k)} = \frac{(-1)(-2)\dots(-k+1)}{(1+x)^k} = (-1)^{k-1} \frac{(k-1)!}{(1+x)^k}$$

II. Derivatives of implicit functions. If y is an implicit function its higher derivatives are found by differentiating the required number of times the equation connecting x and y bearing in mind

that y and all its derivatives are functions of the independent variable.

For instance, the second derivative of the function y specified by the equation

$$x^2 + y^2 = 1$$

is found by differentiating the equation twice. This yields, in succession,

$$x + yy' = 0$$

and

$$(x') + (yy')' = 1 + y'^2 + yy'' = 0$$

whence

$$y' = -\frac{x}{y} \text{ and } y'' = -\frac{1+y'^2}{y} = -\frac{x^2+y^2}{y^3} = -\frac{1}{y^3}$$

In this example the function y can be expressed explicitly in terms of x , and therefore the differentiation can be performed directly; for instance, taking the branch $y = \sqrt{1-x^2}$ we find $y'' = -\frac{1}{\sqrt{(1-x^2)^3}}$ which coincides with the previous result.

In the general case the derivative of any order of an implicit function can always be expressed in terms of the independent variable and the function itself. For instance, on differentiating twice, with respect to x , the equation

$$y^3 - 3axy + x^3 = 0$$

specifying y as an implicit function of x we receive

$$y^2y' - ay - axy' + x^2 = 0$$

and

$$2y(y')^2 + y^2y'' - 2ay' - axy'' + 2x = 0$$

On finding y' from the first relation and substituting it into the second we arrive at an equation from which y'' can be expressed in terms of x and y .

III. Derivatives of functions represented parametrically. In order to find a derivative of higher order of a function specified by parametric equations we differentiate the expression of the preceding derivative considering it a composite function of the independent variable. Let $x = \varphi(t)$, $y = f(t)$ then

$$y' = \frac{dy}{dx} = \frac{f'(t)}{\varphi'(t)}$$

Furthermore,

$$y'' = \frac{d\left(\frac{f'(t)}{\varphi'(t)}\right)}{dx} = \frac{d\left(\frac{f'(t)}{\varphi'(t)}\right)}{dt} \frac{dt}{dx} = \frac{f''(t)\varphi'(t) - f'(t)\varphi''(t)}{\varphi'^2(t)} \frac{dt}{dx}$$

and, since

$$\frac{dt}{dx} = \frac{1}{\varphi'(t)}$$

we arrive at the expression

$$y'' = \frac{f''(t)\varphi'(t) - f'(t)\varphi''(t)}{\varphi'^3(t)}$$

The differentiation of the last relation with respect to x leads to the expression for the third derivative, etc.

Examples. Let us find the second derivative y'' of the function specified by the equations

$$x = \cos t, \quad y = \sin t$$

On differentiating, we obtain

$$y' = \frac{\cos t}{-\sin t} = -\cot t \quad \text{and} \quad y'' = \frac{\frac{d(-\cot t)}{dt}}{-\sin t} = -\frac{1}{\sin^3 t}$$

IV. Differentiating products. Leibniz' rule. The derivative of the n th order of a linear combination of functions $u(x)$ and $v(x)$ is readily expressed in terms of the corresponding derivatives of $u(x)$ and $v(x)$. Namely, if C_1 and C_2 are constants, then

$$(C_1 u + C_2 v)^{(n)} = C_1 u^{(n)} + C_2 v^{(n)}$$

The rule for finding the derivative of the n th order of a product of functions is essentially more complicated. Let u and v be functions of x and

$$y = uv$$

It is required to express $y^{(n)}$ in terms of the functions u , v and their derivatives. Differentiating we find, in succession,

$$\begin{aligned} y' &= u'v + uv' \\ y'' &= u''v + 2u'v' + uv'' \\ y''' &= u'''v + 3u''v' + 3u'v'' + uv''' \end{aligned}$$

Now it is possible to discover the analogy between the expressions of the second and third derivatives and the binomial formulas for, respectively, the second and third powers: the coefficients in the products of the derivatives of the zeroth, first, second and third orders (the *zeroth derivative* coincides, by definition, with the given function) of the functions u and v are equal to the corresponding binomial coefficients in the products of the zeroth, first, second and third powers of u and v entering into the binomial formulas for $(u+v)^2$ and $(u+v)^3$. It turns out that this analogy is of a general character:

Leibniz' Rule. For any n there holds the formula

$$y^{(n)} = (uv)^{(n)} = u^{(n)}v + nu^{(n-1)}v' + \frac{n(n-1)}{2!}u^{(n-2)}v'' + \dots \\ \dots + nu'v^{(n-1)} + uv^{(n)} \quad (*)$$

This formula can be formally obtained if we take Newton's binomial formula for $(u+v)^n$ and then replace the powers of u and v by the derivatives of the corresponding orders of u and v (and put $u^{(0)}=u$, $v^{(0)}=v$).

Here we do not present the general proof of this formula and confine ourselves to considering a characteristic example of its application.

Example. Let us compute the hundredth derivative of the function $y=x^2 \sin x$. We have

$$y^{(100)} = (x^2 \sin x)^{(100)} = (\sin x \cdot x^2)^{(100)} = (\sin x)^{100} x^2 + \\ + 100 (\sin x)^{(99)} (x^2)' + \frac{100 \cdot 99}{2} (\sin x)^{(98)} (x^2)''$$

All the subsequent terms are omitted here since they are identically equal to zero: each of them involves a derivative of order higher than the second of the function x^2 , and $(x^2)^{(k)}=0$ for $k > 2$. Consequently (see I),

$$y^{(100)} = x^2 \sin \left(x + 100 \cdot \frac{\pi}{2} \right) + 200x \sin \left(x + 99 \cdot \frac{\pi}{2} \right) + \\ + 9900 \sin \left(x + 98 \cdot \frac{\pi}{2} \right) = x^2 \sin x - 200x \cos x - 9900 \sin x$$

55. Differentials of Higher Orders. The differential dy of a function $y=f(x)$ depends on two arguments (see Sec. 50) which are the independent variable x and its differential dx . The *differential dx of the independent variable x* is a magnitude independent of x : for any given value of x the values of dx can be chosen quite arbitrarily.

Let us consider $df(x)$ as a function of x (for an arbitrary but fixed value of dx independent of x) and take its differential $d[df(x)]$. If this differential exists we call it the *second differential* or the *differential of the second order* of the function $f(x)$ and denote it by d^2y :

$$d^2y = d(dy)$$

Similarly, the *third differential* or the *differential of the third order* d^3y of a function $y=f(x)$ is the differential of its second differential as a function of x . Now we give the general definition:

Definition. The *differential of the n th order* d^ny of a function $y=f(x)$ is the differential of its $(n-1)$ th differential as a function of x :

$$d^ny = d(d^{(n-1)}y)$$

To unify the terminology, we call the differential $df(x)$ of a function $f(x)$ the *differential of the first order* or the *first differential*.

Let us derive the expression for the differential of a higher order of a function $f(x)$ dependent on the argument x . In all the calculations below the differential dx is understood as an arbitrary but constant quantity independent of x .

For the second differential we obtain

$$d(dy) = (dy)' dx = [f'(x) dx]' dx = f''(x) dx dx$$

that is

$$d^2y = f''(x) dx^2$$

Analogously, for the differential of the n th order we arrive at the formula

$$d^ny = (d^{n-1}y)' dx = [f^{(n-1)}(x) dx^{n-1}]' dx = f^{(n)}(x) dx^n$$

where dx^n is the n th power of dx . Thus,

the differential of the n th order is equal to the product of the n th derivative with respect to the independent variable by the n th power of the differential of the independent variable.

Every differential is an infinitesimal of higher order than all the differentials of lower orders:

$$d^ny = o(d^my) \text{ as } dx \rightarrow 0$$

if $n > m$ and $f^{(m)}(x) \neq 0$.

The sequence of the differentials

$$dy, d^2y, d^3y, \dots, d^ny, \dots$$

is arranged in increasing orders of smallness (as $dx \rightarrow 0$).

If $y = f(x)$ then $dy = f'(x) dx$ irrespective of whether the argument x is an independent variable or a function of another argument (see the invariance property of the first differential in Sec. 51, III). In the general case differentials of higher orders do not possess this property. Indeed, suppose that x is no longer an independent variable, as before, but a function $x = \varphi(t)$ of a new independent variable t . Then dx also becomes a function of t , and therefore it is not allowable to regard dx as a constant when the first differential is differentiated. This leads to a new expression of d^2y different from the above. Namely, computing the differential of dy by applying the differentiation rule for a product we find

$$d^2y = d[f'(x) dx] = d[f'(x)] dx + f'(x) d(dx)$$

But we have

$$d[f'(x)] = f''(x) dx \text{ and } d(dx) = d^2x$$

and therefore

$$d^2y = f''(x) dx^2 + f'(x) d^2x$$

We see that there appears the additional term $f'(x) d^2x$. It vanishes if x is an independent variable:

$$d^2x = (x)'' dx^2 = 0 \cdot dx^2 = 0$$

The expression of the third differential in the case when x depends on t is still more complicated:

$$d^3y = f'''(x) dx^3 + 3f''(x) dx d^2x + f'(x) d^3x$$

Thus, when finding a higher-order differential we should take into account the nature of the argument of the function and distinguish between the cases when it is an independent variable or depends on some other variable.

The formulas for the differentials of the first and higher orders imply the following expressions of the derivatives as ratios of differentials:

$$\begin{aligned} y' &= f'(x) = \frac{dy}{dx} \\ y'' &= f''(x) = \frac{d^2y}{dx^2} \\ &\dots \dots \dots \\ y^{(n)} &= f^{(n)}(x) = \frac{d^ny}{dx^n} \\ &\dots \dots \dots \end{aligned}$$

All these formulas except the first one are only applicable if x is an independent variable.

For the sake of convenience the derivative $\frac{d^ny}{dx^n}$ is sometimes written in the symbolic notation $\frac{d^n}{dx^n} y$. For example, we write $\frac{d^2}{dx^2} (2x^4 - x + 1) = 48x$.

QUESTIONS

1. How are the velocity of motion, the linear density, the heat capacity and the reaction rate defined?

2. What do we call the rate of change of a function?

3. State the definition of the derivative of a given function. Apply this notion to the magnitudes mentioned in Questions 1 and 2.

4. What is the tangent line to a curve at a given point?

5. What is the geometrical meaning of the derivative of a given function $y = f(x)$ in Cartesian coordinates?

6. State the differentiation rules for the results of arithmetical operations and give illustrative examples of their application.

7. How do we differentiate a function of a function and the inverse of a function?

8. Derive the differentiation formulas for all the basic elementary functions.

9. Describe the technique of using the logarithmic derivative.

10. How do we differentiate an implicit function? Give examples.

11. Describe the method of representing functions and curves parametrically. Give examples.

12. How are functions specified parametrically differentiated?

13. State the definition of the angle between two intersecting curves. What is the normal to a curve at a given point?

14. Describe the technique of graphical differentiation.

15. What is the geometrical meaning of the derivative in polar coordinates? Derive the corresponding formula.

16. What do we call the differential of a function? How is the differential of a function expressed in terms of its derivative?

17. Describe the geometrical interpretation of the differential of a given function $y = f(x)$.

18. Enumerate the basic properties of the differential of a function. Describe the invariance property of the form of the first differential.

19. What do we call a differentiable function? What is the necessary condition for the differentiability?

20. Give examples of continuous functions which are nondifferentiable.

21. Write down the formula for an approximation to the value of a function using the differential of the function. Give illustrative examples.

22. Write down the formulas expressing the limiting absolute and relative errors in the computation of the value of a function in terms of a given limiting absolute error of the argument.

23. State the definition of the derivative of the n th order of a given function.

24. How do we compute the derivatives of higher orders of explicit, implicit and parametrically represented functions?

25. State Leibniz' rule for differentiating a product of functions.

26. What is the differential of the n th order of a given function? How is it expressed in terms of the corresponding derivative with respect to the independent variable? Is the form of a higher-order differential invariant?

Chapter IV

APPLICATION OF DIFFERENTIAL CALCULUS TO INVESTIGATION OF BEHAVIOUR OF FUNCTIONS

§ 1. Theorems of Fermat, Rolle, Lagrange and Cauchy

56. Theorems of Fermat and Rolle. We begin with some fundamental theorems of differential calculus.

Fermat's Theorem¹. Let $y=f(x)$ be a continuous function in a closed interval $[x_1, x_2]$ assuming its greatest (or least) value at an interior point ξ of that interval: $x_1 < \xi < x_2$. Then, if the derivative $f'(x)$ of the function $f(x)$ at the point ξ exists it is necessarily equal to zero: $f'(\xi)=0$.

Proof. For definiteness, let the function $f(x)$ assume its greatest value at the point ξ ; this means that

$$f(x) \leq f(\xi) \quad (*)$$

for all $x \in [x_1, x_2]$.² The value of the derivative at the point ξ is

$$f'(\xi) = \lim_{\Delta x \rightarrow 0} \frac{f(\xi + \Delta x) - f(\xi)}{\Delta x}$$

where the increment Δx can be positive or negative since ξ is an interior point of the interval.

Let us consider the ratio

$$\frac{f(\xi + \Delta x) - f(\xi)}{\Delta x}$$

According to condition (*), its numerator cannot be positive: $f(\xi + \Delta x) - f(\xi) \leq 0$. Therefore, this ratio is nonpositive for $\Delta x > 0$:

$$\frac{f(\xi + \Delta x) - f(\xi)}{\Delta x} \leq 0 \text{ if } \Delta x > 0$$

The limit of the ratio, as $\Delta x \rightarrow 0$, $\Delta x > 0$, is the *right-hand derivative* of $f(x)$ at the point ξ (Sec. 52) and is also nonpositive:

¹ P. Fermat (1601-1665), a famous French mathematician.

² We remind the reader that a function continuous in a closed interval necessarily attains its greatest and least values on that interval. In this theorem it is supposed that the point ξ at which the greatest or the least value is attained lies inside the interval, that is, it does not coincide with its end points.

$f'_r(\xi) \leq 0$ (see Sec. 29, II). But if $\Delta x < 0$ the ratio is nonnegative:

$$\frac{f(\xi + \Delta x) - f(\xi)}{\Delta x} \geq 0 \text{ if } \Delta x < 0$$

Consequently, for the left-hand derivative we have $f'_l(\xi) \geq 0$.

By the hypothesis, the function $f(x)$ has the derivative $f'(x)$ at the point ξ , and therefore the right-hand and the left-hand derivatives must coincide. This being only possible if $f'_l(\xi) = f'_r(\xi) = 0$, it follows that $f'(\xi) = 0$, which is what we wished to prove.

Let the reader introduce into this argument the necessary changes for the case when the function attains its *least* value at the point ξ .

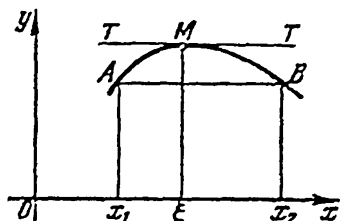


Fig. 64

The geometrical meaning of Fermat's theorem is clearly seen in Fig. 64: the tangent to the graph of a function at its highest (or lowest) point is parallel to the axis of abscissas. If the greatest or the least value of a function is achieved at an end point of the interval of definition (for instance, the function whose graph is

depicted in Fig. 64 assumes its least value at the points $x = x_1$ and $x = x_2$) the tangent at the corresponding point of the graph is not necessarily parallel to the axis of abscissas, that is the derivative at that end point may be nonzero.

It should be noted that there are cases when a function attains its greatest or least value at a point where the derivative does not exist. For example, the function $y = |x|$ assumes its least value (equal to zero) at the point $x = 0$ where its derivative does not exist. The investigation of the behaviour of this function and its graph are given in Sec. 52.

Rolle's theorem is a consequence of Fermat's theorem.

Rolle's Theorem¹. If a function $f(x)$ is continuous in a closed interval $[x_1, x_2]$, is differentiable at all its interior points and assumes equal values at the end points of the interval then there is at least one point $x = \xi$ lying inside that interval at which $f'(\xi) = 0$.

Proof. The end point values of the function being equal ($f(x_1) = f(x_2)$), there are two possible cases here.

(1) The function does not vary within the interval: $f(x) = f(x_1) = f(x_2)$ everywhere, then its derivative is identically equal to zero for all the values of x .

(2) The function varies in the closed interval $[x_1, x_2]$. Since it is continuous it attains its greatest and least values in this

¹ M. Rolle (1652-1719), a French mathematician.

interval (Sec. 35), and at least one of these values is assumed inside the interval $[x_1, x_2]$, for if the greatest and the least values were both assumed at the end points the condition of the theorem would imply that they are equal and the function is constant. In Fig. 64 we see the case when the function attains its *greatest* value inside the interval at an interior point ξ . According to the hypothesis, the derivative exists at all *interior* points of the interval, and hence at the interior point where the function assumes its greatest or least value as well. By Fermat's theorem, the derivative vanishes at that point, which is what we set out to prove.

Rolle's theorem is interpreted geometrically as follows:

On the arc $\bar{AB}$ of the curve $y = f(x)$ (see Fig. 64) connecting the points A and B there is a point lying between the end points A and B of the arc at which the tangent is parallel to the axis of abscissas.

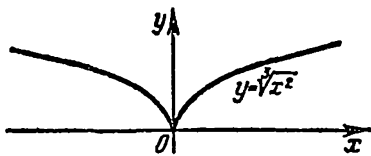


Fig. 65

If $f(x_1) = f(x_2) = 0$ the theorem can be restated as *between every two zeros of a differentiable function there is at least one zero of its derivative.*

M. Rolle himself originally proved this theorem for polynomial functions for this very case.

If the condition that the given function is differentiable at all interior points of the interval is dropped Rolle's theorem no longer holds. For instance, the continuous function $y = \sqrt[3]{x^2}$ assumes equal values at the end points of the interval $[-1, +1]$ ($y_{x=-1} = y_{x=1} = 1$), and at the same time its derivative $y' = \frac{2}{3\sqrt[3]{x}}$ does

not vanish anywhere. In this example the conditions of the theorem are violated at the point $x=0$ lying inside the interval $(-1, +1)$ since the derivative does not exist at this point¹. The graph of this function is shown in Fig. 65. At the point $(0, 0)$ it has the tangent perpendicular to the axis of abscissas, and the derivative does not exist at that point; the graph consists of two branches approaching the origin in such a way that they touch that common tangent at the point $(0, 0)$. Such a point of a curve is called a *cusp* or a *spinode*.

Rolle's theorem can also be inapplicable when the function is not continuous in the closed interval $[x_1, x_2]$. For example, this is the case for the function y specified in the interval $[0, 1]$ by

¹ This follows from the argument in footnote¹ on p. 162.

the conditions

$$y = \begin{cases} x & \text{for } 0 \leq x < 1 \\ 0 & \text{for } x = 1 \end{cases}$$

This function satisfies all the requirements of Rolle's theorem except the condition of continuity: it is discontinuous at the right end point $x=1$ of the interval $[0, 1]$.

It is clearly seen that the assertion of the theorem does not hold: the derivative is equal to 1 at all the points of the interval $(0, 1)$, and therefore there is no point at which it turns into zero.

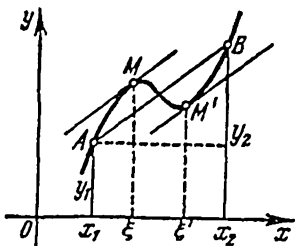


Fig. 66

If the function itself is everywhere continuous and the derivative does not exist at the end points of the interval but does exist inside it the theorem remains true. An example of this kind is the function $y = \sqrt{1-x^2}$ considered in the interval $[-1, 1]$ where it is continuous while its derivative does not exist for $x = \pm 1$. Let the reader check the validity of Rolle's theorem for this case.

57. Lagrange's¹ Theorem. Let us come back to the geometrical interpretation of Rolle's theorem. The chord subtending the arc AB (Fig. 64) being parallel to the axis of abscissas, the tangent at the point M is parallel to the chord. It is clear that this property is retained when the chord subtending an arc is not parallel to the axis of abscissas. Suppose that AB is an arc of a continuous line $y=f(x)$ having a nonvertical tangent at each point (Fig. 66). Then *there always is at least one point M on the arc AB at which the tangent is parallel to the chord subtending this arc.*

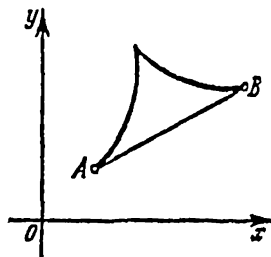


Fig. 67

If the above conditions on the arc AB do not hold such a point does not necessarily exist (see Fig. 67).

Now, let the arc AB be specified by the equation $y=f(x)$ where $f(x)$ is a differentiable function, the abscissas of the points A and B being, respectively, x_1 and x_2 (Fig. 66). Since the slope of the chord AB is equal to $\frac{f(x_2)-f(x_1)}{x_2-x_1}$ and that of the tangent line at the point M (there are two such points in Fig. 66: M and M') is equal to $f'(\xi)$ where ξ is the abscissa of the point M ,

¹ J.L. Lagrange (1736-1813), an outstanding French mathematician and astronomer.

the condition of the parallelism of the tangent and the chord implies the equality of their slopes. We thus arrive at the following proposition:

Lagrange's Theorem. If $f(x)$ is a continuous function in a closed interval $[x_1, x_2]$ differentiable at all its interior points there is at least one point $x = \xi$ inside the interval such that

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi)$$

We now proceed to prove this theorem *analytically*. To this end, let us take the auxiliary function

$$F(x) = f(x) - \lambda x$$

where λ is the slope of the straight line N_1N_2 parallel to the chord M_1M_2 and passing through the origin (Fig. 68). Thus λ is equal to the slope of the chord M_1M_2 :

$$\lambda = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

The function $F(x)$ satisfies all the requirements of Rolle's theorem in the interval $[x_1, x_2]$: it is continuous for $x \in [x_1, x_2]$ and differentiable for $x \in (x_1, x_2)$ (since the functions $f(x)$ and λx possess these properties) and assumes equal values at the end points of the interval:

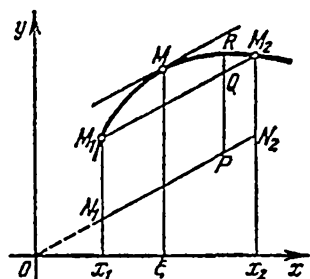


Fig. 68

$$F(x_1) = f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} x_1 = \frac{x_2 f(x_1) - x_1 f(x_2)}{x_2 - x_1}$$

and

$$F(x_2) = f(x_2) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} x_2 = \frac{x_2 f(x_1) - x_1 f(x_2)}{x_2 - x_1}$$

The geometrical meaning of the auxiliary function $F(x)$ is quite evident. The ordinate of the variable point of its graph is represented by the line segment PR (Fig. 68) which is equal to the sum of the constant segment

$$PQ = N_1M_1 = N_2M_2$$

and the segment QR . Consequently, the chord connecting the end points of the arc specified by the new function is parallel to Ox .

By Rolle's theorem, there exists a point $\xi \in (x_1, x_2)$ such that $F'(\xi) = 0$. But $F'(x) = f'(x) - \lambda$, and therefore $f'(\xi) - \lambda = 0$ whence

$$\lambda = \frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi)$$

which is what we had to prove.

It follows from Lagrange's theorem that

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \quad x_1 < \xi < x_2$$

This relation is known as the *formula of finite increments*. It states that

the increment of a differentiable function on an interval is equal to the product of the derivative of the function at an intermediate point by the increment of the independent variable.

The formula of finite increments makes it possible to find the exact expression for the increment of a function in terms of the increment of the argument and the value of the derivative at an interior point of the interval. It has a significant theoretical importance and lies in the foundation of the proofs of a number of important theorems which will be discussed later on.

58*. Cauchy's¹ Theorem. Cauchy's theorem is a generalization of Lagrange's theorem. We begin with its analytical formulation and proof.

Cauchy's Theorem. If $f(x)$ and $\varphi(x)$ are continuous functions in a closed interval $[x_1, x_2]$ differentiable at all its interior points and if the derivative $\varphi'(x)$ does not turn into zero in the interior of that interval there exists at least one point $x = \xi$ lying inside the interval such that

$$\frac{f(x_2) - f(x_1)}{\varphi(x_2) - \varphi(x_1)} = \frac{f'(\xi)}{\varphi'(\xi)}$$

It should be noted that $\varphi(x_2) - \varphi(x_1) \neq 0$ since, if otherwise, there would exist, according to Rolle's theorem, an interior point of the interval at which the derivative $\varphi'(x)$ would vanish, which contradicts the hypothesis.

Proof. Let us take an auxiliary function similar to the one used in the proof of Lagrange's theorem: denoting the ratio $\frac{f(x_2) - f(x_1)}{\varphi(x_2) - \varphi(x_1)}$ by λ we put

$$F(x) = f(x) - \lambda \varphi(x)$$

The auxiliary function $F(x)$ satisfies all the conditions of Rolle's theorem, because its continuity and differentiability follow from the similar properties of the functions $f(x)$ and $\varphi(x)$ and the coincidence of its values at the end points of the interval is readily checked:

$$\begin{aligned} F(x_1) &= f(x_1) - \lambda \varphi(x_1) & -\frac{f(x_1)}{\varphi(x_1)} \varphi(x_1) &= \frac{f(x_1)\varphi(x_2) - f(x_2)\varphi(x_1)}{\varphi(x_2) - \varphi(x_1)} \\ \text{and} & & & \\ F(x_2) &= f(x_2) - \lambda \varphi(x_2) & -\frac{f(x_2)}{\varphi(x_2)} \varphi(x_2) &= \frac{f(x_1)\varphi(x_2) - f(x_2)\varphi(x_1)}{\varphi(x_2) - \varphi(x_1)} \end{aligned}$$

Consequently, by Rolle's theorem, there is a point ξ in the interval (x_1, x_2) such that $F'(\xi) = 0$. But $F'(x) = f'(x) - \lambda\varphi'(x)$, and therefore $f'(\xi) - \lambda\varphi'(\xi) = 0$ whence

$$\lambda = \frac{f(x_2) - f(x_1)}{\varphi(x_2) - \varphi(x_1)} = \frac{f'(\xi)}{\varphi'(\xi)}$$

The theorem has been proved.

In the particular case $\varphi(x) = x$ Cauchy's theorem gives the same result as Lagrange's theorem.

Note that Cauchy's theorem cannot be proved by a simple term-by-term division of the relations expressing Lagrange's theorem for the functions f and φ since the intermediate values ξ of the argument do not necessarily coincide in these relations.

Let us show that Cauchy's theorem can be given the same geometrical interpretation as Lagrange's theorem. For the sake of convenience, let us denote the independent variable by t and consider the functions $y = f(t)$ and $x = \varphi(t)$ as representing parametric equations of a plane curve in the xy -plane. As the parameter t runs through the corresponding interval $[t_1, t_2]$ the variable point (x, y) describes a curve whose initial and terminal points are, respectively, $[\varphi(t_1), f(t_1)]$ and $[\varphi(t_2), f(t_2)]$. The slope of the chord connecting these points is equal to the ratio $\frac{f(t_2) - f(t_1)}{\varphi(t_2) - \varphi(t_1)}$.

The derivative of y (regarded as a parametrically represented function of x) with respect to x is $\frac{dy}{dx} = \frac{f'(t)}{\varphi'(t)}$. Consequently, the formula

$$\frac{f(t_2) - f(t_1)}{\varphi(t_2) - \varphi(t_1)} = \frac{f'(\xi)}{\varphi'(\xi)} \quad (t_1 < \xi < t_2)$$

again expresses the equality of the slopes of the chord subtending an arc and of the tangent line drawn to that arc at an intermediate point.

§ 2. Investigating Functions with the Aid of First and Second Derivatives

59. Testing Functions for Monotonicity¹. The application of the differential calculus to the investigation of functions is based on a simple relationship between the behaviour of a function and the properties of its derivatives and, particularly, of the first derivative.

¹ It is advisable to reread Sec. 13 before proceeding to study the subject matter of this section.

Theorem (Necessary Conditions for Monotonicity).

(1) If a differentiable function $f(x)$ increases in an interval its derivative $f'(x)$ is nonnegative: $f'(x) \geq 0$.

(2) If a differentiable function $f(x)$ decreases in an interval its derivative $f'(x)$ is nonpositive: $f'(x) \leq 0$.

(3) If a differentiable function $f(x)$ does not vary in an interval (i.e. is equal to a constant) its derivative is identically equal to zero: $f'(x) = 0$.

Proof. (1) Let $f(x)$ be a differentiable and increasing function in a given interval; this means that $f(x + \Delta x) > f(x)$ if $\Delta x > 0$ and $f(x + \Delta x) < f(x)$ if $\Delta x < 0$ for any x belonging to this interval. It follows that the ratio

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x} \quad (*)$$

is positive irrespective of the sign of Δx , and therefore its limit as $\Delta x \rightarrow 0$, that is the derivative of $f(x)$, is a positive number or zero.

(2) If $f(x)$ decreases the ratio $\frac{\Delta y}{\Delta x}$ is negative and its limit is a negative number or zero, that is $f'(x) \leq 0$.

In (1) and (2) it is supposed that the function $f(x)$ possesses a derivative at every point. Let the reader draw a graph representing a function which is monotone and continuous but has no derivative at separate points (see Sec. 52).

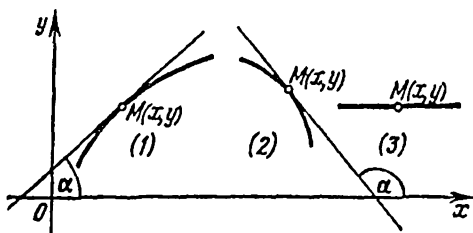


Fig. 69

(3) If $f(x)$ is constant its derivative is known to be equal to zero: $f'(x) = 0$.

The geometrical meaning of this theorem is elucidated by Fig. 69. When the variable point $M(x, y)$ traces the graph of the function from left to right, i.e. as the abscissa increases, it moves upward, along Oy , if the function increases; then, as a rule, its tangent line forms an acute angle α with Ox , and $\tan \alpha > 0$. If the variable point $M(x, y)$ moves downward then, as a rule, the angle α between the tangent and the axis of abscissas is obtuse, and $\tan \alpha < 0$.

It should be stressed that the derivative of a monotone function may vanish at some separate points. For instance, the cubic function $y = x^3$ increases throughout Ox but its derivative $y' = 3x^2$ turns into zero at the point $x = 0$ although at all the other points it is positive. Geometrically, this means that the tangent to the graph of a monotone function may be parallel to Ox at some points.

Thus, in an interval of monotonicity of a function its derivative cannot change sign to the opposite.

This proposition allows us to judge upon the sign of the derivative of a monotone differentiable function in a given interval by its increase or decrease in this interval. But when we begin to investigate a given function its behaviour is usually not known, and therefore it is much more important to establish the converse of the above proposition which enables us to study the *character of the variation* of a function in a given interval by reducing the problem to the simpler question of determining the *sign of its derivative*.

Theorem (Sufficient Condition for Monotonicity).

Let $f(x)$ be a continuous function in a closed interval $[a, b]$ possessing the derivative $f'(x)$ at all the interior points of that interval. Then:

(1) If the derivative $f'(x)$ is everywhere positive inside the interval the function $f(x)$ increases in the interval $[a, b]$.

(2) If the derivative $f'(x)$ is everywhere negative inside the interval the function $f(x)$ decreases in the interval $[a, b]$.

(3) If the derivative $f'(x)$ is everywhere equal to zero inside the interval the function $f(x)$ does not vary in the interval $[a, b]$ (i.e. it is constant).

Proof. Let us choose in the given closed interval $[a, b]$ two arbitrary points x_1 and x_2 such that $a \leq x_1 < x_2 \leq b$ and apply the formula of finite increments to the interval $[x_1, x_2]$:

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \quad x_1 < \xi < x_2$$

Since $x_1 < x_2$, the difference $x_2 - x_1$ is positive, and the sign of the difference $f(x_2) - f(x_1)$ is completely specified by the sign of the derivative $f'(\xi)$. If the derivative $f'(x)$ is *everywhere* positive we have $f'(\xi) > 0$, and, consequently, $f(x_2) - f(x_1) > 0$, that is

$$f(x_2) > f(x_1) \text{ for } x_2 > x_1$$

which means that the function $f(x)$ increases¹.

If the derivative is *everywhere* negative we have $f'(\xi) < 0$, and, consequently, $f(x_2) < f(x_1)$ for $x_2 > x_1$, that is, the function decreases.

Finally, if the derivative is *everywhere* equal to zero we have $f'(\xi) = 0$, and, consequently, $f(x_2) = f(x_1)$ for any x_1 and x_2 , which means that the function $f(x)$ is constant.

This proposition provides a simple *analytical test for monotonicity* of a function in a given interval. For example, take the function $y = \frac{2}{3}x^3 - 2x^2 - 6x + 3$. To check that in the interval

¹ The argument in this proof shows that at the points a and b the derivative may not exist. It should also be noted that a and b may, respectively, be equal to $-\infty$ and $+\infty$.

$-1 < x < 3$ it decreases it is sufficient to verify that its derivative $y' = 2x^2 - 4x - 6$ is negative for $-1 < x < 3$. The latter does in fact take place since $y' = 2(x+1)(x-3)$, and the factor $(x+1)$ is positive for all the values of x in this interval while the factor $(x-3)$ is negative.

Geometrically, it appears evident that if the derivative of a function takes zero values at some separate points but retains constant sign at all the other points the function is also monotone in the given interval. For instance, the function $y = x - \sin x$ increases in every interval since its derivative $y' = 1 - \cos x$ is everywhere positive except the points $x = 2k\pi$ at which it vanishes.

For functions increasing or decreasing in the broad sense (see page 41) we can readily state a test providing necessary and sufficient conditions for monotonicity:

A continuous function $f(x)$ defined in an interval $[a, b]$ and differentiable at all its interior points is nondecreasing (nonincreasing) if and only if the derivative $f'(x)$ is nonnegative: $f'(x) \geq 0$ (respectively, nonpositive: $f'(x) \leq 0$).

60. Extrema of Functions.

1. A value of x separating an interval of increase of a function from an interval of decrease (see Fig. 70) is an important element

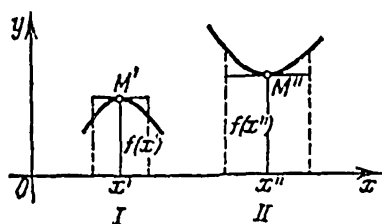


Fig. 70

of the behaviour of the function. At such a point the function $f(x)$ changes the character of its behaviour: as the independent variable passes through that point (from left to right, as we agreed), the function $f(x)$ passes from increase to decrease (see graph I in Fig. 70) or from decrease to increase (graph II in Fig. 70). In the former case

(see the point x' in Fig. 70 separating an interval of increase from an interval of decrease) there exists a neighbourhood of the point x' such that the value $f(x')$ is *greater* than all the other values of the function $f(x)$ within this neighbourhood. In the latter case (see the point x'' in Fig. 70 separating an interval of decrease from an interval of increase) there is a neighbourhood of the point x'' such that the value $f(x'')$ is *less* than all the other values of the function $f(x)$ in this neighbourhood. Thus $f(x')$ and $f(x'')$ are, respectively, the *greatest* and the *least* values of the function in these neighbourhoods. Now let us state the general definition.

Definition. A point x_0 is called a *point of maximum* of the function $f(x)$ if $f(x_0)$ is the greatest value of the function $f(x)$ in a neighbourhood of the point x_0 .

Similarly,

A point x_0 is called a *point of minimum* if $f(x_0)$ is the least value of the function $f(x)$ in a neighbourhood of the point x_0 .

The points of maximum and minimum are called *points of extremum* of the function. Accordingly, we say that $f(x)$ has a *maximum* or a *minimum* (or, generally, an *extremum*) at such a point. We also say that $f(x_0)$ is a *maximum* or a *minimum* or, generally, an *extremal* value of the function $f(x)$.

When speaking of a maximum or a minimum at a point x_0 we usually mean a *strict* extremum. In such a case, there exists a neighbourhood of the point of extremum x_0 for all whose points (except the point x_0 itself) there holds the strict inequality

$$f(x) < f(x_0) \text{ for the point of maximum}$$

or the strict inequality

$$f(x) > f(x_0) \text{ for the point of minimum}$$

If the signs $>$ and $<$ in these inequalities are replaced, respectively, by $\geq$ and $\leq$ the point x_0 is called a *point of nonstrict extremum*.

A function considered in an interval can have several extrema, and it may turn out that a minimum value of the function is greater than a maximum value (see Fig. 71). Therefore a *minimum* or a *maximum* of a function is spoken of as a *relative extremum* (*relative maximum* or *relative minimum*) when the values of the function are compared within a small neighbourhood of the corresponding point. In contrast to it, when speaking of the *greatest* or the *least* value of a function for the whole interval in which it is considered we use the term *absolute extremum* (*absolute maximum* or *absolute minimum*).

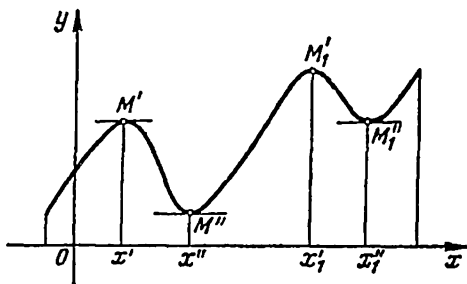


Fig. 71

On the graph of a function, to the points of extremum there correspond "tops" and "cavities" (see the points M' , M'' , M'_1 and M''_1 in Fig. 71).

Usually we deal with functions having only a finite number of extrema in a given finite interval; as a rule, in what follows we shall consider functions of that kind.

An example of a function of a different kind is $y = x \sin \frac{1}{x}$; in every interval containing the point 0 it is continuous (if we

additionally put $y_{x=0}=0$) and has an infinite number of points of extremum. Its graph can be constructed schematically if we resort to the hint concerning the construction of the graph of the function $y = \sin \frac{1}{x}$ on page 102.

Now we proceed to establish a *necessary condition (test)* for an extremum.

Necessary Condition for Extremum. If a function $f(x)$ has an extremum at a point x_0 its derivative at that point either is equal to zero or does not exist.

Indeed, if the function attains a maximum at the point x_0 its value at this point is the *greatest* in a neighbourhood of the point x_0 . By Fermat's theorem, at the interior points of the interval of definition where a differential function attains its greatest value the derivative of the function vanishes. The argument is completely similar in the case of a minimum.

Geometrically, this means that *the tangent to the graph of a function is parallel to Ox at its "tops" and "cavities"* (see Fig. 71).

A function can also have extrema at some of the points where it is nondifferentiable. Examples of that kind are demonstrated in Fig. 72. At the points M_1 and M_2 the graph shown in the

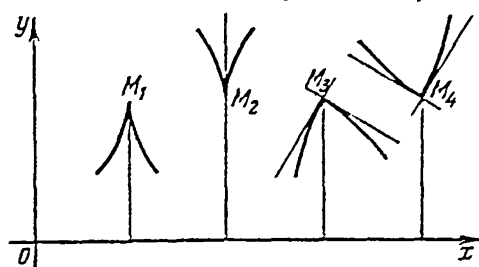


Fig. 72

figure has tangents perpendicular to the axis of abscissas. As was mentioned, a point of this type is called a cusp or a spinode. A feature of such a point is that, as x approaches the abscissa of the point, the derivative $f'(x)$ tends to $+\infty$ on one side of the point and to $-\infty$ on the other side. At the points M_3 and

M_4 the slope of the tangent changes in jump-like manner, and the graph does not have a definite direction of the tangent at these points; such points are termed *corner points* of the curve. For the values of x equal to the abscissas of corner points the right-hand and the left-hand derivatives of the function $f(x)$ (see Sec. 52) have different values.

It should be noted that *the necessary condition for extremum is not sufficient*: the fact that the derivative at a given point turns into zero (or does not exist) does not necessarily imply that this point is a point of extremum. For example, the derivative of the function $y = x^3$ is $y' = 3x^2$; it vanishes at the point $x = 0$ but this point is not a point of extremum of the function (see Fig. 73a). Similarly, the function $y = \sqrt[3]{x}$ (for which

$y' = \frac{1}{3\sqrt[3]{x^2}}, x \neq 0$) has no derivative at the point $x=0$, but

Fig. 73b shows that in this example as well the point $x=0$ is not a point of extremum. It is readily seen that in both examples the point $x=0$ does not separate intervals of monotonicity of different senses: both on the right and on the left of this point the derivative has the same sign.

To find out whether a given point where the derivative turns into zero or does not exist is a point of extremum we should resort to *sufficient conditions (tests)* for extremum to which we proceed now.

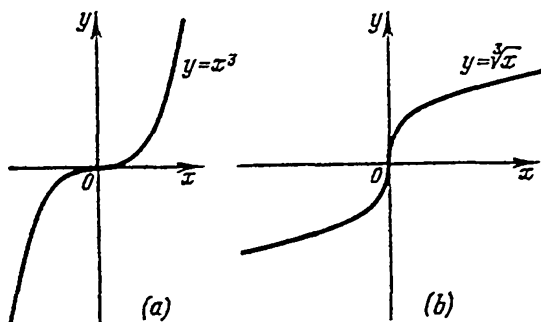


Fig. 73

Sufficient Condition for Extremum in Terms of the First Derivative. Let $f(x)$ be a continuous function in a neighbourhood of a point x_0 possessing a derivative everywhere in this neighbourhood except possibly at the point x_0 itself; then:

If the derivative $f'(x)$ is positive for $x < x_0$ and negative for $x > x_0$ the point x_0 is a point of maximum.

If the derivative $f'(x)$ is negative for $x < x_0$ and positive for $x > x_0$ the point x_0 is a point of minimum.

Here, according to the necessary condition for extremum, it is supposed that at the point x_0 itself the derivative either is equal to zero or does not exist.

Proof. Let $f'(x) > 0$ for $x < x_0$. This means that on the left of the point x_0 there is an interval of increase of the function $f(x)$ adjoining the point x_0 . If $f'(x) < 0$ for $x > x_0$ then on the right of the point x_0 there is an interval of decrease of the function adjoining the point x_0 . Consequently, x_0 is a point of maximum.

The other case when the derivative changes its sign from $-$ to $+$ as x passes through the point x_0 from left to right is investigated quite similarly.

Thus, briefly: if, as x passes through the point x_0 (from left to right), the derivative changes sign the point x_0 is a point of extremum.

If the derivative changes sign from $+$ to $-$ there is a maximum at the point x_0 ; if it changes sign from $-$ to $+$ there is a minimum at the point x_0 .

It is clear that if the derivative $f'(x)$ does not change sign as x passes through the point x_0 there is no extremum at the point x_0 . For instance, let $f'(x) > 0$ both for $x < x_0$ and for $x > x_0$. Then the function increases both on the left and on the right of the point x_0 , and therefore it cannot assume at this point its greatest value compared with all the values in a neighbourhood of the point x_0 . We have just discussed examples of this kind, namely, the behaviour of the functions $y = x^3$ and $y = \sqrt[3]{x}$ in the vicinity of the point $x = 0$ (see Fig. 73 and the explanation given above).

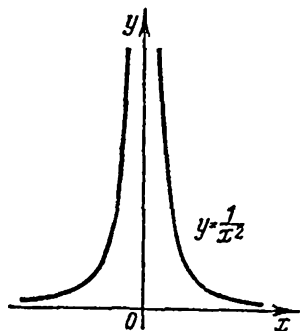


Fig. 74

Remark. It should be noted that if it is only known that the derivative changes sign at a point, it is impossible to judge upon the existence of an extremum because it is necessary to know additionally that the function is continuous at that point itself. For instance, take the function $y = \frac{1}{x^2}$. Its derivative $y' = -\frac{2}{x^3}$ changes sign as x passes through the point $x = 0$: $y' > 0$ for $x < 0$ and $y' < 0$ for $x > 0$; consequently, the function increases on the

left of $x = 0$ and decreases on the right of $x = 0$. At the same time, $x = 0$ is not a point of maximum of the function since it has an infinite discontinuity at that point. This situation is clarified in Fig. 74.

II. Scheme for investigating functions by means of the sufficient condition for extremum in terms of the first derivative. Let us indicate a sequence of operations for finding the points of extremum and the intervals of monotonicity of a function on the basis of the test proved above. Consider a function $y = f(x)$ continuous in a given (finite or infinite) interval and possessing a derivative everywhere except possibly at some separate points. For the sake of simplicity, we shall suppose that if these points are deleted the derivative $f'(x)$ is continuous in the remaining intervals.

(1) First of all, we should find the points of the interval, where the given function is considered, at which the derivative is equal to zero (these are called *stationary points*), that is the real roots of the equation $f'(x) = 0$, and also the points at which the derivative does not exist. Let us arrange all these points in an increasing sequence and denote them $x_1, x_2, \dots, x_n$:

$$x_1 < x_2 < \dots < x_n$$

We have thus determined the set of all points at which the function $f(x)$ may have extrema. They are sometimes termed *critical points* (this set contains both the stationary points and the points at which the derivative does not exist).

We shall additionally suppose that if the interval in which the function is investigated is finite it only contains a finite number of critical points. If the interval is infinite there may be an infinitude of critical points. For example, the function $y = \sin x$ when considered throughout Ox has infinitely many stationary points which are the roots of the equation $y' = \cos x = 0$, i.e. the points $x = \frac{\pi}{2} + k\pi$ where $k = 0, \pm 1, \pm 2, \dots$. It is apparent that the function $y = \sin x$ has an extremum at each of these points.

(2) Next we split the whole interval $[a, b]$ in which the function is considered, by means of the points x_i , into the *subintervals* (a, x_1) , (x_1, x_2) , $\dots$, (x_{n-1}, x_n) , (x_n, b) . In each subinterval the derivative has constant sign. For, by the hypothesis, the derivative is continuous in each subinterval, and hence, if it changed the sign, $f'(x)$ would turn into zero at some new point (see Sec. 35), which is impossible since all such points have already been found (they serve as the end points of the subintervals).

Thus, these subintervals are the intervals of monotonicity of the function.

(3) Now we must determine the *sign of the derivative* in each subinterval, for which it is sufficient to determine it at an arbitrary point of every subinterval. The sign of the derivative specifies the character of variation of the function in each interval of monotonicity, that is its increase or decrease. The specification of the change of sign of the derivative as x passes through the boundaries of the intervals of monotonicity, that is through the points x_i , indicates which of these points are points of maximum and which points of minimum. It may also turn out that some of the points x_i are not points of extremum. This is the case when the function has the same character of monotonicity in two adjoining subintervals (x_{i-1}, x_i) and (x_i, x_{i+1}) separated by the point x_i , i.e. when the derivative has the same sign in them. Then both subintervals form an entire interval of monotonicity of the function. In this case x_i is not a point of extremum of the function (for instance, for the function $y = x^3$ the point $x = 0$ belongs to this type).

(4) The substitution of the critical values $x = x_i$ into $f(x)$ yields the corresponding values of the function:

$$f(x_1), f(x_2), \dots, f(x_n)$$

As has been already mentioned, not all of them are necessarily extremal.

If the values $f(x_1), f(x_2), \dots, f(x_n)$ have been computed and the end-point values $f(a)$ and $f(b)$ have also been found the character of variation of the function becomes clear without investigating the sign of the derivative. For, since we know that the function has no extrema within each of the subintervals $(a, x_1), (x_1, x_2), \dots, (x_n, b)$ and thus is monotone on them, the comparison of the values of the function at the end points of each subinterval indicates directly whether the function increases or decreases on that subinterval.

It depends on the concrete circumstances which of the techniques should be used. It is sometimes more convenient to consider the sign of the derivative in the subintervals, but since usually we are interested in the values of the function at the critical points the second approach may prove preferable.

It is advisable to compile a table of the following type:

x	a	$a < x < x_1$	x_1	$x_1 < x < x_2$	x_2	$x_2 < x < b$	b
y y'	$f(a)$	y increases +	$f(x_1)$ 0	y decreases —	$f(x_2)$ 0	y decreases —	$f(b)$

For the sake of simplicity this table is formed for the illustrative case when there are only two critical points in the interval $[a, b]$.

The sign of the derivative in each subinterval can be determined by computing the value of the function $f'(x)$ at an arbitrary point of every subinterval. If the derivative is represented as a product of a number of factors it is sufficient to determine the signs of these factors without computing their values since these signs specify the sign of the derivative (see Example 2 below).

Such a table makes it possible to draw a sketch of the graph of the function.

III. Examples of investigating functions.

(1) Basic elementary functions. Although the behaviour of the basic elementary functions was already described with the aid of their graphs (see Chapter I) the application of the first derivative makes it easy to investigate analytically the simplest properties of these functions.

Power function: $y = x^\alpha, x > 0$. Its derivative $y' = \alpha x^{\alpha-1}$ is positive for $x > 0$ if $\alpha > 0$ and negative if $\alpha < 0$. Consequently, on the positive half-axis Ox the function either everywhere increases (if $\alpha > 0$) or everywhere decreases (if $\alpha < 0$).

Exponential function: $y = a^x$. The sign of its derivative $y' = a^x \ln a$ coincides with that of $\ln a$ since the function a^x itself is everywhere positive. Hence, $y' > 0$ if $a > 1$ and $y' < 0$ if $a < 1$. Thus.

the exponential function is monotone throughout Ox : it increases in the former case and decreases in the latter.

Logarithmic function: $y = \ln x$. The function is only defined for $x > 0$. For these values of x its derivative $y' = \frac{1}{x}$ is positive, and hence the function $y = \ln x$ is everywhere increasing.

Similarly, resorting to the first derivative we can readily determine the character of variation of all the other basic elementary functions.

(2) Let us consider the function

$$y = 3x^3 + 4.5x^2 - 4x + 1$$

It possesses a continuous derivative throughout Ox . On computing the derivative

$$y' = 9x^2 + 9x - 4 = (3x - 1)(3x + 4)$$

we see that it is equal to zero for $x = -\frac{4}{3}$ and for $x = \frac{1}{3}$. Both factors on the right being negative for $x < -\frac{4}{3}$, the derivative is positive for these values of x , and thus the function increases. In the interval $-\frac{4}{3} < x < \frac{1}{3}$ the derivative is negative, and the function decreases. Finally, for $x > \frac{1}{3}$ the derivative is again positive, and the function increases. Consequently, $x = -\frac{4}{3}$ is a point of maximum and $x = \frac{1}{3}$ a point of minimum, and there are three intervals of monotonicity: the function increases in the interval $(-\infty, -\frac{4}{3}]$, decreases in the interval $[-\frac{4}{3}, \frac{1}{3}]$ and again increases in the interval $[\frac{1}{3}, +\infty)$.

On computing the values of the function at the points of extremum we can compile a table characterizing the behaviour of the function. When compiling it we should take into account that

$$\lim_{x \rightarrow +\infty} y = \lim_{x \rightarrow +\infty} x^3 \left(3 + \frac{4.5}{x} - \frac{4}{x^2} + \frac{1}{x^3} \right) = +\infty$$

and, similarly, $\lim_{x \rightarrow -\infty} y = -\infty$.

x	$-\infty$	$-\infty < x < -\frac{4}{3}$	$-\frac{4}{3}$	$-\frac{4}{3} < x < \frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3} < x < +\infty$	$+\infty$
y	$-\infty$	y increases	$7\frac{2}{9}$	y decreases	$\frac{5}{18}$	y increases	$+\infty$
y'		+	$\max$ 0	-	$\min$ 0	+	

x	-2	$-2 < x < -\frac{2}{3}$	$-\frac{2}{3}$	$-\frac{2}{3} < x < 0$	0	$0 < x < 0.5$	0.5
y, y'	≈ 0.21	y increases +	$y_{\max} \approx 0.39$ 0	y decreases —	$y_{\min} = 0$ ∞	y increases +	≈ 1.03

characterizing the behaviour of the function we see that in the first interval the derivative is positive and the function increases, in the second interval the derivative is negative and the function decreases and in the third interval the derivative is again positive and the function increases. By the way, this becomes clear from the comparison of the values of the function assumed at the critical points with the end-point values:

$$y_{x=-2} \approx 0.21, y_{x=-\frac{2}{3}} = \sqrt[3]{\frac{4}{9}} e^{-\frac{2}{3}} \approx 0.39, y_{x=0} = 0, y_{x=0.5} \approx 1.03$$

The greatest value of the function on the interval $[-2, 0.5]$ is approximately equal to 1.03 and it is attained at the right end point of the interval while the least value is equal to 0 and is reached at the point $x=0$.

Later on (see page 203) we shall continue the investigation of this function to obtain some additional information about its behaviour.

(4) The methods of the investigation of the behaviour of functions can be applied to proving *inequalities*. For instance, let us prove the inequalities

$$\frac{2}{\pi} x < \sin x < x \text{ for } 0 < x < \frac{\pi}{2}$$

To this end, we introduce the function y specified by the conditions

$$y = \begin{cases} \frac{\sin x}{x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0 \end{cases}$$

Its derivative

$$y' = \frac{\cos x}{x^2} (x - \tan x)$$

being negative in the interval $(0, \frac{\pi}{2})$ (since $x < \tan x$), y is a decreasing function in that interval. Hence, for $0 < x < \frac{\pi}{2}$, the values of the expression $\frac{\sin x}{x}$ are less than the value 1 of the function y reached at the point $x=0$ and exceed the value

$\frac{2}{\pi} = \frac{\sin \frac{\pi}{2}}{\frac{\pi}{2}}$ of that function attained at the point $x = \frac{\pi}{2}$:

$$\frac{2}{\pi} < \frac{\sin x}{x} < 1$$

This implies the desired inequalities.

61. Investigating Functions with the Aid of Second Derivatives. The Greatest and the Least Values of a Function.

1. With the aid of the second derivative $f''(x)$ of the function $f(x)$ under investigation it is possible to establish another sufficient test for extremum which sometimes proves simpler and more convenient than the one in the foregoing section. In what follows we assume that in a neighbourhood of a given point x_0 the function $f(x)$ itself and its first and second derivatives are continuous.

Sufficient Condition for Extremum in Terms of the Second Derivative. If the first derivative vanishes at the point x_0 ($f'(x_0) = 0$) while the second derivative is different from zero ($f''(x_0) \neq 0$), x_0 is a point of extremum. Furthermore,

if $f''(x_0) < 0$ the point x_0 is a point of maximum,

if $f''(x_0) > 0$ the point x_0 is a point of minimum.

Proof. Let $f'(x_0) = 0$ and $f''(x_0) > 0$. By the hypothesis, the second derivative is continuous, and therefore its sign is retained in a neighbourhood of the point x_0 . It follows that the function $f'(x)$ increases in this neighbourhood because its derivative $(f'(x))' = f''(x)$ is positive. Since $f'(x_0) = 0$, the derivative $f'(x)$ assumes values less than $f'(x_0) = 0$ on the left of the point x_0 and is therefore negative: $f'(x) < 0$ for $x < x_0$. Similarly, on the right of the point x_0 its values are greater than $f'(x_0) = 0$, that is positive: $f'(x) > 0$ for $x > x_0$. Hence, as x passes through the point x_0 from left to right, the function $f'(x)$ changes sign from $-$ to $+$, and therefore, according to the foregoing test for extremum (in terms of the first derivative), x_0 is a point of minimum of the function $f(x)$.

An analogous argument shows that if $f''(x_0) < 0$ the function $f'(x)$ decreases and changes its sign from $+$ to $-$ as x passes through the point x_0 , that is the point x_0 is a point of maximum of the function $f(x)$.

If both $f'(x_0) = 0$ and $f''(x_0) = 0$ the latter test is inapplicable, and one should resort to the former. For instance, the first and the second derivatives of the function $y = x^4$ turn into zero at the point $x = 0$. Therefore, the second test is inapplicable while the first one indicates that there is a minimum at that point since the derivative $y' = 4x^3$ changes sign from $-$ to $+$ as x passes through the origin. At the same time, the function $y = x^3$ whose first and second derivatives also vanish at the point $x = 0$ has no

extremum at the origin; indeed, its first derivative $y' = 3x^2$ does not change sign as x passes through the point $x = 0$.

But when applicable, the second test proves extremely convenient since it does not require the determination of the sign of the function $f'(x)$ at points different from the point at which the given function is tested for extremum, and this makes it possible to judge upon the existence of the extremum by the sign of the function $f''(x)$ at the same point.

Examples. (1) Let $y = ax^2 + bx + c$. Since the derivative $y' = 2ax + b$ is equal to zero at the point $x = -\frac{b}{2a}$, and $y'' = 2a$, the function y has a maximum at that point $x = -\frac{b}{2a}$ if $a < 0$ and a minimum if $a > 0$.

This example demonstrates the importance of theoretical investigations for practical problems of mathematics. It shows that the study of a concrete function is simplified and that the amount of calculations is reduced as we use more sophisticated theoretical means of mathematical analysis. For instance, the investigation of the quadratic function presented in Chapter I required a special argument and rather lengthy calculations since we could not use the differential calculus while the present consideration is quite simple.

(2) Let us come back to Example 2 on page 191 in which the function $y = 3x^3 + 4.5x^2 - 4x + 1$ was considered. Its derivative $y' = 9x^2 + 9x - 4$ is equal to zero at the points $x = -\frac{4}{3}$ and $x = \frac{1}{3}$. The second derivative has an extremely simple expression: $y'' = 18x + 9$. It is readily seen that $y''_{x=-\frac{4}{3}} < 0$, that is $x = -\frac{4}{3}$ is a point of maximum, and that $y''_{x=\frac{1}{3}} > 0$, which means that $x = \frac{1}{3}$ is a point of minimum.

II. The greatest and the least values of a function. After the extremal values of a function $y = f(x)$ considered on a closed interval $[a, b]$ have been found we easily determine the *greatest value* M and the *least value* m of this function in that interval. For this purpose the values of the function at the points of extremum should be compared with each other and with the end-point values. For the greatest (the least) value of the function on the interval $[a, b]$ is either one of its maximum (minimum) values or an end-point value. Some of the possible cases are shown in Fig. 77.

It is also clear that when a function $y = f(x)$ is monotone in a closed interval $[a, b]$ its greatest value is $f(b)$ and the least

value is $f(a)$ if the function increases and, conversely, the greatest value is $f(b)$ if the function decreases.

It often occurs that a given function has *only one* point of extremum in an interval. In this case the value of the function at that point is the greatest on the interval in the case of maximum and the least in the case of minimum.

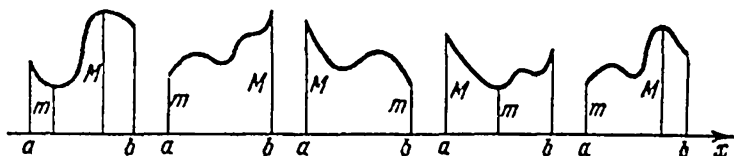


Fig. 77

III. Some problems involving the determination of the greatest and the least values. Suppose that there are two magnitudes connected by a functional relationship, and it is required to find the value of one of them (belonging to an interval that can be finite or infinite) for which the other magnitude assumes its *least* or *greatest* value among all the possible values. To solve such a

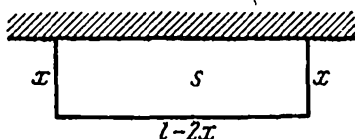


Fig. 78

problem we must first find the expression of the function describing the relationship between the magnitudes in question and then determine the least or the greatest value of this function on the given interval.

Examples. (1) Let us determine the least length l of the fence enclosing a rectangular lot of land with given area s adjoining a wall (see Fig. 78).

Denoting by x one of the sides of the rectangle we readily obtain

$$s = x(l - 2x)$$

whence

$$l = 2x + \frac{s}{x}$$

The problem now reduces to finding the least value of this function as x ranges from 0 to ∞ . Since $l \rightarrow \infty$ both for $x \rightarrow 0$ and for $x \rightarrow \infty$, the least value is among the minima of the function

in the interval $(0, \infty)$. Let us find the derivative:

$$l' = 2 - \frac{s}{x^2}$$

It follows that in the interval in question there is only one stationary point $x = \sqrt{\frac{s}{2}}$ at which the function has a minimum since the second derivative $l'' = \frac{2s}{x^3}$ is positive at that point. Hence, the minimum value

$$l_{\min} = 2\sqrt{\frac{s}{2}} + \frac{s}{\sqrt{\frac{s}{2}}} = 2\sqrt{2s}$$

serves simultaneously as the least value of the function on the entire interval $(0, \infty)$. Consequently, the length of any fence enclosing a rectangular lot of land with a given area s adjoining a wall cannot be less than $2\sqrt{2s}$, and it is equal to this value only when the smaller side of the rectangle (which is equal to $\sqrt{\frac{s}{2}} = \frac{1}{2}\sqrt{2s}$) is half the greater side (which is equal to $2\sqrt{2s} - 2 \cdot \frac{1}{2}\sqrt{2s} = \sqrt{2s}$). Thus, in these circumstances the most economical fence is the one whose greater side is twice the smaller one.

(2) Consider a free fall of a rain drop (air resistance is neglected) of initial mass m_0 . Suppose that it evaporates uniformly so that the decrement of its mass is proportional to time of fall t , the proportionality factor being k . It is required to find how long it takes the kinetic energy of the drop to achieve its greatest value.

The kinetic energy a as function of time of fall t is given by the expression

$$a = \frac{mv^2}{2} = \frac{(m_0 - kt)(gt)^2}{2}$$

where g is the acceleration of gravity. The problem thus reduces to finding the greatest value of this function for $t > 0$.

We have

$$\frac{da}{dt} = -\frac{k}{2}(gt)^2 + (m_0 - kt)g^2t = g^2t \left(-\frac{3}{2}kt + m_0 \right)$$

and it is clear that $\frac{da}{dt} = 0$ for $t = 0$ and for $t = \frac{2}{3} \frac{m_0}{k}$. Furthermore, $\frac{da}{dt} > 0$ for $0 < t < \frac{2}{3} \frac{m_0}{k}$ and $\frac{da}{dt} < 0$ for $t > \frac{2}{3} \frac{m_0}{k}$. It follows that

for $t = \frac{2m_0}{3k}$ the kinetic energy a attains a maximum which coincides with its greatest value since there are no other maxima. Thus, the greatest value is

$$a_{\max} = \frac{\left(m_0 - k \frac{2m_0}{3k}\right) \left(g \frac{2m_0}{3k}\right)^2}{2} = \frac{2}{27} \frac{g^2 m_0^3}{k^2}$$

If the drop falls on the ground in time $T \leq \frac{2m_0}{3k}$ the kinetic energy increases during the whole time of motion, and its greatest value is reached at the moment the drop falls on the ground.

62. Convexity and Concavity of a Curve. Points of Inflection.

1. Considering an arc of a curve of the second order, that is of an ellipse (whose particular case is a circle), a parabola or a hyperbola, we readily observe that such an arc lies entirely on one side of the tangent line drawn at any point of the arc. On the other hand, for the graph of the function $y = \sin x$ (the sinusoid) this property holds for the arcs corresponding to the intervals $[-\pi, 0]$ and $[0, \pi]$ of the axis of abscissas but does not hold for the arc corresponding to the entire interval $[-\pi, \pi]$. For if we draw the tangent to the sinusoid at the origin the left and the right parts of this curve are on different sides of the tangent.

These considerations lead us to the following definition:

Definition. An arc of a curve is said to be *convex* if it lies entirely on one side of the tangent drawn through any point of the arc.

It is of course supposed here that the tangent can be drawn at each point of the arc, that is there are neither corner points nor cusps on it.

According to this definition, any arc of a second-order algebraic curve and the arcs of the sinusoid corresponding to the intervals $[-\pi, 0]$ and $[0, \pi]$ are convex whereas the arc of the sinusoid corresponding to the whole interval $[-\pi, \pi]$ is not. The definition of convexity is demonstrated in Fig. 79: the arc AB in Fig. 79a is convex while the arc CD in Fig. 79b is not since, for instance, it does not lie on one side of the tangent line passing through the point M .

This definition of convexity of an arc is coherent with the notion of a convex polygon and a convex broken line defined in elementary geometry: *a polygon is said to be convex if it is placed entirely on one side of every straight line in which its any side lies*, and *a nonclosed broken line is considered convex if the polygon appearing as its end points are joined by a line segment is convex*.

Now we proceed to test for convexity a curve which is the graph of a continuous function $y = f(x)$ whose first and second

derivatives are also continuous. If such a curve is convex we distinguish between the cases of an arc *convex upward* and an arc *convex downward* (see, respectively, Fig. 80, I and II). Namely, an arc of a curve specified in the xy -plane by an equation $y=f(x)$ is said to be *convex upward* if it entirely lies *below* the tangent

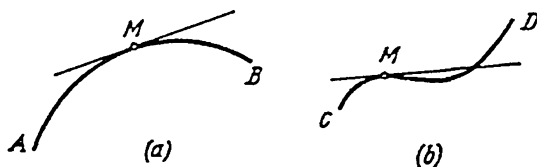


Fig. 79

drawn through each point of the arc. Similarly, it is said to be *convex downward* if it entirely lies *above* its tangent at every point.

We also speak about *concavity* of curves: an arc convex upward (downward) is said to be *concave downward* (*upward*).

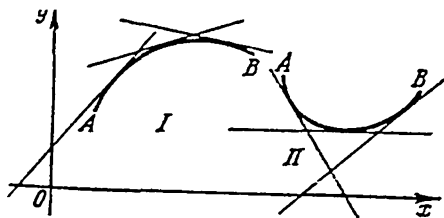


Fig. 80

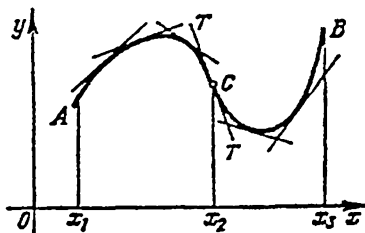


Fig. 81

Finally, a convex upward (i.e. concave downward) arc is sometimes briefly called *convex* while a concave upward (i.e. convex downward) arc is simply referred to as *concave*¹. In what follows we shall use the latter variant of the terminology.

An important characteristic of a curve are the points separating its convex and concave arcs; for instance, in Fig. 81 the point C separates the convex arc AC from the concave arc CB.

We shall always suppose that any arc of a curve under consideration corresponding to a finite interval of the axis of abscissas only contains a finite number of such points².

¹ It should be noted that the notion of a convex function introduced in modern mathematics corresponds to a function $y=f(x)$ whose graph is convex downward.

² An example of a curve of a different kind is the graph of the function $y=x \sin \frac{1}{x}$ (see page 185).

Definition. A point of a curve separating its convex arc from a concave arc is termed a *point of inflection*.

At a point of inflection the tangent intersects the curve; in the vicinity of such a point the curve lies on both sides of its tangent drawn through that point.

Up to now we spoke of the graph of a function $y=f(x)$ possessing a derivative, that is of a curve at whose every point there is a tangent not perpendicular to the axis of abscissas. But it should be noted that a point separating a convex arc of a curve from a concave one may be such that the tangent at that point is perpendicular to the axis of abscissas or such that the tangent does not exist. The first of these possibilities is demonstrated by the behaviour of the graph of the function $y=\sqrt[3]{x}$ in the vicinity of the origin (see Fig. 73b on p. 187); in such a case we speak of a *point of inflection with vertical tangent*.

Let the reader draw a curve with a corner point separating its convex and concave arcs; we shall not include corner points of this kind into the class of points of inflection. There may of course exist a corner point at which the character of convexity of the curve does not change (see Fig. 67 on p. 178).

II. Testing a curve for convexity and concavity. As before, we consider a curve $y=f(x)$ where the function $f(x)$ is continuous together with its first and second derivatives $f'(x)$ and $f''(x)$. There is a simple relationship between the properties of the second derivative $f''(x)$ and the convexity or concavity of the curve $y=f(x)$ which we shall establish without a rigorous proof by resorting only to some simple geometrical considerations. These considerations are based on the following proposition.

On an interval of decrease of the first derivative the graph of the function is convex and on the interval of increase of the first derivative it is concave.

For, if the derivative $f'(x)$ decreases, the slope of the tangent line to the curve also decreases as the variable point of the graph traces it from left to right; then the tangent line turns as if it "pressed down" the curve preventing it from lifting above the tangent. Therefore the arc must lie below its every tangent and thus is convex. This situation is clearly demonstrated by the example of the arc AC of the curve depicted in Fig. 81. On the contrary, if the derivative $f'(x)$ increases the slope of the tangent increases as well, and the tangent turns as if it "lifted up" the curve preventing it from descending below the tangent. Therefore the corresponding arc must lie above its every tangent and thus is concave (see the arc CB in Fig. 81).

Now we can take advantage of the basic theorem establishing the connection between the character of variation of a function and the sign of its derivative; as a function under consideration

we take $f'(x)$ whose derivative is $(f'(x))' = f''(x)$. If $f''(x) > 0$ the derivative $f'(x)$ increases and if $f''(x) < 0$ it decreases. We thus arrive at the following theorem:

Theorem. If the second derivative $f''(x)$ is everywhere negative within an interval the arc of the curve $y = f(x)$ corresponding to that interval is convex. If the second derivative $f''(x)$ is everywhere positive in an interval the corresponding arc of the curve $y = f(x)$ is concave.

As at the end of Sec. 59, it should be noted that if the second derivative $f''(x)$ has constant sign, for instance, if it is positive, everywhere except at some separate points where it vanishes, the function $f'(x)$ remains increasing, and the corresponding arc of the graph of the function $y = f(x)$ is concave. For instance, the second derivative of the function $y = x^4$ (whose graph is a parabola of the fourth order) is $y'' = 12x^2$; this derivative is everywhere positive except at the point $x = 0$ where it turns into zero whereas the graph of the parabola is everywhere concave.

The theorem stated here elucidates the test for extremum of a function based on investigating its second derivative. Indeed, if $f'(x_0) = 0$ and $f''(x_0) < 0$ the graph of the function has a horizontal tangent line at the point x_0 and is *convex* in the vicinity of that point; therefore x_0 is a point of maximum. Similarly, if $f'(x_0) = 0$ and $f''(x_0) > 0$ the tangent to the graph of the function $y = f(x)$ at the point x_0 is horizontal, and the graph itself is *concave*; therefore the point x_0 is a point of minimum.

III. Tests for the determination of points of inflection. A point of inflection of the graph of a function is a boundary between its convex and concave arcs. It follows from the above argument that the abscissa of a point of inflection (see the point x_0 in Fig. 81) separates two adjoining intervals of monotonicity of the first derivative $f'(x)$ and is therefore a *point of extremum of the first derivative*. Consequently, the application of the necessary condition for extremum implies that

If x_0 is the abscissa of a point of inflection then $f''(x_0) = 0$.

This proposition provides a *necessary test (condition)* for the point of inflection.

But by far not every root of the equation $f''(x) = 0$ is the abscissa of a point of inflection. For example, the first and the second derivatives of the function $y = x^4$ are $y' = 4x^3$ and $y'' = 12x^2$. Although $y''_{x=0} = 0$, the point $(0, 0)$ is the vertex of the parabola $y = x^4$ and does not serve as its point of inflection.

In order to find out whether or not a value x_0 of the argument for which the second derivative vanishes is the abscissa of a point of inflection we should investigate the variation of the sign of the second derivative as x passes through this point, that is to find out whether the second derivative changes sign. This argument

leads to the following *sufficient* test for the point of inflection.

If the second derivative $f''(x)$ changes sign as x passes through x_0 (from left to right) then x_0 is the abscissa of a point of inflection. If it changes sign from $-$ to $+$ there is an interval of convexity on the left of the point x_0 and an interval of concavity on the right of it, and, conversely, if it changes sign from $+$ to $-$, an interval of convexity follows an interval of concavity as x passes through x_0 .

In conclusion we note once again that when determining points of inflection and intervals of convexity and concavity we use the same rules as in the case of the determination of points of extremum but apply these rules not to the given function itself but to its first derivative. Our investigation shows that *to an interval of increase of the first derivative there corresponds an interval of concavity of the graph of the function and to an interval of decrease of the first derivative an interval of convexity; accordingly, the abscissa of a point of extremum of the first derivative is that of a point of inflection of the graph of the function.*

Examples. (1) The second derivative of the function $y = x^3 + x$ is $y'' = 20x^3$, and $y''_{x=0} = 0$. We have $y'' < 0$ for $x < 0$ and $y'' > 0$ for $x > 0$. Consequently, the origin is a point of inflection of the graph of the given function, the interval of convexity lying on the left of it and the interval of concavity on the right.

(2) Let us take the function $y = x^5 + 5x^4$. Here $y'' = 20x^3(x+3)$, and $y'' = 0$ for $x = -3$ and for $x = 0$. As x passes through the point $x = -3$ the second derivative changes sign, and thus $x = -3$ is the abscissa of a point of inflection. When x passes through the point $x = 0$ the second derivative retains constant sign, and therefore the origin is not a point of inflection; the graph of the given function is concave on both sides of the origin.

In our foregoing investigation we supposed that the function $f(x)$ in question was twice differentiable throughout the interval in question. If this condition is violated it is necessary to investigate $f'(x)$ and $f''(x)$ in the vicinity of those separate points at which the derivatives do not exist. At these points as well the character of convexity of the graph of the function may change. An example of this kind is the graph of the function $y = \sqrt[3]{x}$ (see Fig. 73b) for which the point $(0, 0)$ is a point of inflection.

In most typical cases the points of extremum $a, b, c, \dots$ and the abscissas of points of inflection $\alpha, \beta, \gamma, \dots$ alternate (see Fig. 71).

Note that there is a difference in the terminology: a point of extremum of a function is a point lying on the *axis* along which the independent variable is reckoned while a point of inflection is a point lying on the *graph* itself. This is accounted for by the fact that an extremum is a relative notion dependent on the system of coordinates chosen since at the point of a curve corresponding

to an extremum the tangent line is parallel to the axis of abscissas, and if another coordinate system is taken whose axes are turned through an angle relative to the former axes this property of the tangent is not retained. At the same time, a feature of a point of inflection is that the tangent intersects the curve at that point, and this property is characteristic of the curve itself and is invariant with respect to any change of the coordinate system; the property of a point to be a point of inflection of a curve is retained under all the displacements of the curve (as rigid body) in the plane.

IV. Examples. (1) The graphs of the basic elementary functions can readily be tested for convexity and concavity.

For instance, the graph of the power function $y = x^\alpha$ is concave on the positive half-axis Ox for $\alpha < 0$ and for $\alpha > 1$ and convex for $0 < \alpha < 1$. Indeed, in the first two cases the second derivative $y'' = \alpha(\alpha - 1)x^{\alpha-2}$ is positive and in the third case negative.

The graph of the exponential function $y = a^x$ is concave for any $a > 0$ throughout Ox since $y'' = a^x \ln^2 a > 0$.

The logarithmic curve $y = \log_a x$ is convex for $a > 1$ and concave for $a < 1$ because the second derivative $y'' = -\frac{1}{\ln a} \frac{1}{x^2}$ is negative in the former case and positive in the latter.

Since we have

$$(\sin x)'' = -\sin x$$

the convex arcs of the sinusoid lie above the axis of abscissa and the concave arcs below it; the points of inflection of the sinusoid are those at which it cuts Ox .

(2) Let us dwell in more detail on the function

$$y = \sqrt[3]{x^2} e^x, \quad -2 \leq x \leq 0.5$$

(see Sec. 60, III, Example 3). The investigation of the intervals of monotonicity and points of extremum already performed shows that this function has one point of maximum $x = -\frac{2}{3}$ and one point of minimum $x = 0$ within the interval $[-2, 0.5]$. Now we proceed to determine the intervals of convexity and concavity and the points of inflection. For this purpose we compute the second derivative:

$$y'' = \left(\frac{2+3x}{3\sqrt[3]{x}} \right)' e^x + \frac{2+3x}{3\sqrt[3]{x}} (e^x)' = \frac{9x^2 + 12x - 2}{9x\sqrt[3]{x}} e^x$$

The factors e^x and $x\sqrt[3]{x} = \sqrt[3]{x^4}$ being positive, the sign of y'' coincides with that of the quadratic trinomial $9x^2 + 12x - 2$. This trinomial has zeros at the points $x_1 = \frac{-\sqrt{6+2}}{3} \approx -1.48$ and $x_2 = \frac{\sqrt{6-2}}{3} \approx 0.15$. It is apparent that $y'' > 0$ in the interval

$(-2, x_1)$, $y'' < 0$ in the interval $(x_1, 0)$, $y'' < 0$ in the interval $(0, x_2)$ and $y'' > 0$ in the interval $(x_2, 0.5)$.¹ Consequently, the graph of the function is concave in the interval $(-2, x_1)$, convex in the intervals $(x_1, 0)$ and $(0, x_2)$, and concave in the interval $(x_2, 0.5)$ (see Fig. 82).

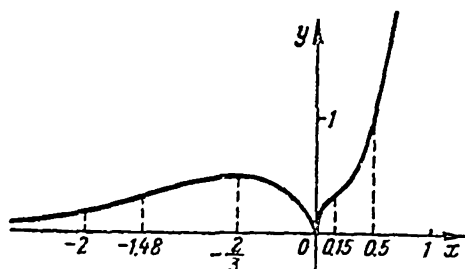


Fig. 82

To construct the graph of the function more accurately let us determine the ordinates of the points of inflection; the substitution of x_1 and x_2 into the expression of the function yields $y_{x \approx -1.48} \approx 0.30$ and $y_{x \approx 0.15} \approx 0.32$. We see that the graph in Fig. 82 is considerably distinct from that in

Fig. 76 which was constructed without taking into account convexity and concavity of the curve and its points of inflection. For further improvement of the graph it is advisable to find, in addition, the slopes of the tangents at the points of inflection. For this purpose it is necessary to substitute x_1 and x_2 into the expression of the first derivative; on calculating we obtain

$$y'_{x \approx -1.48} \approx 0.16, \quad y'_{x \approx 0.15} \approx 1.80$$

It should also be noted that outside the interval $[-2, 0.5]$ we have considered the function has neither points of extremum nor points of inflection.

§ 3. L'Hospital's Rule. General Scheme for Investigating Functions

63. L'Hospital's² Rule.

1. Basic cases. In addition to the rules of passing to the limit discussed in Chapter II we shall present here one more rule known as *L'Hospital's rule* which is extremely simple and convenient. In what follows we shall also apply it to the investigation of the behaviour of functions. This rule can be stated as

L'Hospital's Theorem. Let $f(x)$ and $\varphi(x)$ be two functions tending simultaneously to zero or to infinity as $x \rightarrow x_0$ (or as $x \rightarrow \infty$). If the ratio of their derivatives has a limit (finite or

¹ It is impossible to consider the whole interval $[x_1, x_2]$ without splitting it since it contains the point $x=0$ at which the derivatives do not exist.

² G. F. A. L'Hospital (1661-1704), a French mathematician, a contemporary of Leibniz and Newton. He was the first to publish a book on differential calculus (in 1696).

infinite) the ratio of the functions also possesses a limit which is equal to the limit of the ratio of the derivatives:

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{\varphi'(x)}$$

We shall not give the general proof of this theorem and shall only confine ourselves to some simple considerations and examples elucidating the essence of the proposition¹.

To begin with, let us prove that if the functions $f(x)$ and $\varphi(x)$ tend to zero as $x \rightarrow x_0$ and if their derivatives at the point x_0 exist and $\varphi'(x_0) \neq 0$ then

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \frac{f'(x_0)}{\varphi'(x_0)} \quad (*)$$

It is supposed here that the functions $f(x)$ and $\varphi(x)$ are defined and continuous in a neighbourhood of the point x_0 and that the denominator $\varphi(x)$ does not vanish in this neighbourhood except at the point x_0 itself.

We have $f(x_0) = \varphi(x_0) = 0$, and therefore

$$\frac{f(x)}{\varphi(x)} = \frac{f(x) - f(x_0)}{\varphi(x) - \varphi(x_0)} = \frac{\frac{f(x) - f(x_0)}{x - x_0}}{\frac{\varphi(x) - \varphi(x_0)}{x - x_0}}$$

Passing to the limit as $x \rightarrow x_0$ and using the theorem on the limit of a quotient we obtain

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \frac{\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}}{\lim_{x \rightarrow x_0} \frac{\varphi(x) - \varphi(x_0)}{x - x_0}}$$

whence, by the definition of the derivative,

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \frac{f'(x_0)}{\varphi'(x_0)}$$

L'Hospital's rule is applicable to one-sided limits as well.

Examples.

$$(1) \lim_{x \rightarrow 0} \frac{\sin x}{x} = \frac{(\sin x)'_{x=0}}{1} = 1$$

$$(2) \lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \frac{\left(\frac{1}{1+x}\right)_{x=0}}{1} = 1$$

¹ In the statement of this theorem we have not stipulated some additional conditions on functions $f(x)$ and $\varphi(x)$ guaranteeing the applicability of Cauchy's theorem to them (see the proof of formula (**)) given below; as a rule, in practical examples these conditions are fulfilled.

We see how easily the limits that we found in the foregoing sections can now be determined but it should be noted that here we use the derivatives of the functions $\sin x$ and $\ln(1+x)$ which, in their turn, are obtained on the basis of the computation of the limits of the ratios $\frac{\sin x}{x}$ and $\frac{\ln(1+x)}{x}$ as $x \rightarrow 0$.

$$(3) \lim_{x \rightarrow 2} \frac{x^2 - 5x + 6}{x^2 - 3x + 2} = \frac{2 \cdot 2 - 5}{2 \cdot 2 - 3} = -1$$

$$(4) \lim_{x \rightarrow \pi} \frac{\sin 5x}{\sin 2x} = \frac{5 \cos 5\pi}{2 \cos 2\pi} = -\frac{5}{2}$$

It may turn out that the derivative $\varphi'(x_0)$ is equal to zero at the point x_0 . Then of course formula (*) is inapplicable, and we should resort to the general form of L'Hospital's rule:

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{\varphi'(x)} \quad (**)$$

The proof of formula (**) is more complicated than that of formula (*). It is based on *Cauchy's theorem* (Sec. 58), and the reader who skipped this theorem may omit the proof given below and proceed directly to the examples.

Let us again write down the ratio of the functions in the form

$$\frac{f(x)}{\varphi(x)} = \frac{f(x) - f(x_0)}{\varphi(x) - \varphi(x_0)}$$

and apply to it Cauchy's theorem (to this end we should additionally assume that the derivatives $f'(x)$ and $\varphi'(x)$ exist in a neighbourhood of the point x_0 and that $\varphi'(x)$ does not turn into zero anywhere except at the point x_0 itself). Cauchy's theorem implies

$$\frac{f(x)}{\varphi(x)} = \frac{f'(\xi)}{\varphi'(\xi)}$$

where ξ is a point lying in the interval between x_0 and x . If now $x \rightarrow x_0$ the point ξ contained between x_0 and x also tends to the point x_0 . By the hypothesis, the limit of the ratio of the derivatives exists, and hence we obtain

$$\lim_{x \rightarrow x_0} \frac{f(x)}{\varphi(x)} = \lim_{\xi \rightarrow x_0} \frac{f'(\xi)}{\varphi'(\xi)}$$

On replacing the letter ξ on the right by x we arrive at formula (**).

If the existence of the limit of the ratio of the derivatives is not stipulated formula (**) may be inapplicable although the limit of the ratio of the functions themselves may nevertheless exist (examples of this kind are given below). The matter is that if the limit $\lim_{x \rightarrow x_0} \frac{f'(x)}{\varphi'(x)}$ exists as x tends to x_0 in an arbitrary manner

then of course the limit $\lim_{\xi \rightarrow x_0} \frac{f'(\xi)}{\varphi'(\xi)}$ exists as well, but the converse is *not true* since ξ may assume not all but only some specific values when $x \rightarrow x_0$. This situation is analogous to the fact that although the limit $\lim_{x \rightarrow 0} \sin \frac{1}{x}$ does not exist (for x tending to 0 in an arbitrary way), the limit $\lim_{n \rightarrow \infty} \sin \frac{1}{\left(\frac{1}{\pi n}\right)} = \lim_{n \rightarrow 0} \sin \pi n = 0$ exists ($x = \frac{1}{\pi n} \rightarrow 0$ as $n \rightarrow \infty$).

(5) $\lim_{x \rightarrow 0} \frac{x - x \cos x}{x - \sin x} = \lim_{x \rightarrow 0} \frac{1 - \cos x + x \sin x}{1 - \cos x}$. In the limit on the right the numerator and the denominator, in their turn, also tend to zero. In such cases L'Hospital's rule can be used repeatedly as many times as necessary:

$$\lim_{x \rightarrow 0} \frac{1 - \cos x + x \sin x}{1 - \cos x} = \lim_{x \rightarrow 0} \frac{2 \sin x + x \cos x}{\sin x} = \lim_{x \rightarrow 0} \frac{3 \cos x - x \sin x}{\cos x} = 3$$

Formula (**) also remains valid when the ratio of the derivatives tends to infinity: in this case the ratio of the functions tends to infinity as well. This is demonstrated by the following example.

$$(6) \lim_{x \rightarrow 0} \frac{1 - \cos x}{x - \sin x} = \lim_{x \rightarrow 0} \frac{\sin x}{1 - \cos x} = \lim_{x \rightarrow 0} \frac{\cos x}{\sin x} = \infty$$

When L'Hospital's rule is used repeatedly it is advisable to perform beforehand all the possible simplifications of the given expression, for instance, to cancel by common factors and to use the limits already known.

$$(7) \lim_{x \rightarrow 0} \frac{\arctan x - \arcsin x}{\tan x - \sin x} = \lim_{x \rightarrow 0} \frac{\frac{1}{1+x^2} - \frac{1}{\sqrt{1-x^2}}}{\frac{1}{\cos^2 x} - \cos x} =$$

$$= \lim_{x \rightarrow 0} \frac{\sqrt{1-x^2} - 1 - x^2}{1 - \cos x} \cdot \frac{\cos^2 x}{(1 + \cos x + \cos^2 x)(1 + x^2) \sqrt{1-x^2}}$$

The limit of the second factor is clearly equal to $\frac{1}{3}$; on applying L'Hospital's rule to the first factor we finally receive

$$\frac{1}{3} \lim_{x \rightarrow 0} \frac{\frac{-x}{\sqrt{1-x^2}} - 2x}{\sin x} = -\frac{1}{3} \lim_{x \rightarrow 0} \frac{x}{\sin x} \left(\frac{1}{\sqrt{1-x^2}} + 2 \right) = -1$$

Let us prove that formula (**) remains valid for $x \rightarrow \infty$ on condition that the functions $f(x)$ and $\varphi(x)$ are defined and diffe-

rentiable for sufficiently large values of $|x|$. Indeed, the substitution $x = \frac{1}{z}$ reduces the problem to the case already investigated since $x \rightarrow \infty$ implies $z \rightarrow 0$ and vice versa. Thus, we have

$$\lim_{x \rightarrow \infty} \frac{f(x)}{\varphi(x)} = \lim_{z \rightarrow 0} \frac{f\left(\frac{1}{z}\right)}{\varphi\left(\frac{1}{z}\right)} = \lim_{z \rightarrow 0} \frac{f'\left(\frac{1}{z}\right)\left(-\frac{1}{z^2}\right)}{\varphi'\left(\frac{1}{z}\right)\left(-\frac{1}{z^2}\right)}$$

On cancelling by $\left(-\frac{1}{z^2}\right)$ and again replacing $\frac{1}{z}$ by x we obtain

$$\lim_{x \rightarrow \infty} \frac{f(x)}{\varphi(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{\varphi'(x)}$$

It should be noted that the rule is applicable in all cases when x tends to infinity, that is for $x \rightarrow +\infty$, for $x \rightarrow -\infty$ and for $x \rightarrow \infty$ (see Sec. 25).

If for $x \rightarrow x_0$ (or for $x \rightarrow \infty$) both functions $f(x)$ and $\varphi(x)$ simultaneously tend to infinity formula (**) also remains valid but the proof becomes more sophisticated and we do not treat it here.

$$(8) \lim_{x \rightarrow +\infty} \frac{\ln x}{x^\alpha} = \lim_{x \rightarrow +\infty} \frac{\frac{1}{x}}{\alpha x^{\alpha-1}} = \lim_{x \rightarrow +\infty} \frac{1}{\alpha x^\alpha} = 0 \text{ for } \alpha > 0$$

$$(9) \lim_{x \rightarrow +\infty} \frac{a^x}{x^n} = \lim_{x \rightarrow +\infty} \frac{a^x \ln a}{n x^{n-1}} = \dots = \lim_{x \rightarrow +\infty} \frac{a^x (\ln a)^n}{n!} = \infty$$

if n is a positive integer and $a > 1$. By the way, for $a > 1$ this limit is also equal to ∞ for any n .

Examples 8 and 9 indicate that, as $x \rightarrow +\infty$, the logarithmic function $\ln x$ increases slower and the exponential function a^x faster than any power function x^α (with positive α).

In conclusion we give an example in which the ratio of functions has a limit whereas the ratio of their derivatives does not tend to any limit; we have

$$\lim_{x \rightarrow \infty} \frac{x + \sin x}{x} = \lim_{x \rightarrow \infty} \left(1 + \frac{\sin x}{x}\right) = 2$$

but at the same time the function

$$\frac{(x + \sin x)'}{(x)'} = \frac{1 + \cos x}{1}$$

constantly oscillates assuming the values lying between 0 and 2 as $x \rightarrow \infty$ and hence has no limit. This example does not contradict L'Hospital's rule; simply, the latter is inapplicable to this case.

II. **Some other cases.** In the problems considered above we dealt with limits of ratios $\frac{f(x)}{\varphi(x)}$ when $f(x) \rightarrow 0$ and $\varphi(x) \rightarrow 0$. Condi-

onally, such limits are denoted by the symbol $\frac{0}{0}$ whose meaning is clear. Similarly, the limit of such a ratio as $f(x) \rightarrow \infty$ and $\varphi(x) \rightarrow \infty$ is denoted $\frac{\infty}{\infty}$.

L'Hospital's rule often makes it possible to determine limits in some cases different from $\frac{0}{0}$ and $\frac{\infty}{\infty}$. Here we shall consider limits conditionally denoted as (1) $0 \cdot \infty$, (2) $\infty - \infty$, (3) 1^∞ , (4) ∞ and (5) 0^0 . These expressions, i.e. $0 \cdot \infty$, $\infty - \infty$, 1^∞ , ∞^0 , 0^0 , and also $\frac{0}{0}$ and $\frac{\infty}{\infty}$ only serve for a symbolic indication of a characteristic peculiarity of the limit of a function in question. Namely, they show that the function whose limit we are interested in (as the argument x varies in a certain way, i.e., $x \rightarrow x_0$ or $x \rightarrow \infty$) is, respectively:

(1) the product of a function tending to zero by a function tending to infinity;

(2) the difference of two functions simultaneously approaching plus infinity;

(3) a function tending to unity raised to a power which is a function tending to infinity;

(4) a function tending to infinity raised to a power which tends to zero;

(5) a function tending to zero raised to an infinitesimal power¹.

We shall not dwell on the detailed theoretical analysis of the techniques used for evaluating such limits; they all are based on the reduction to the two basic cases $\frac{0}{0}$ or $\frac{\infty}{\infty}$, and the examples below will show the reader how this reduction is performed practically. It should only be noted that in Cases (3), (4) and (5) we first take the logarithm of the given expression, then find the limit of that logarithm and, finally, determine from the latter the limit of the original function, which is allowable because of the continuity of the logarithmic function.

Examples. (1) The limit $A = \lim_{x \rightarrow 0} x^n \ln x$ ($n > 0$) belongs to the case $0 \cdot \infty$. The transformation to the form

$$A = \lim_{x \rightarrow 0} \frac{\ln x}{\frac{1}{x^n}}$$

reduces it to the case $\frac{\infty}{\infty}$. Finally, the application of L'Hospital's

¹ The solution of problems of finding limits of the types $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0 \cdot \infty$, $\infty - \infty$, 1^∞ , ∞^0 and 0^0 is referred to as the *evaluation of indeterminate forms*.

l'Hôpital's rule results in

$$A = \lim_{x \rightarrow 0} \frac{\frac{1}{x}}{-\frac{n}{x^{n+1}}} = \lim_{x \rightarrow 0} \left(-\frac{x^n}{n} \right) = 0$$

(2) The limit $A = \lim_{x \rightarrow 1} \left(\frac{x}{x-1} - \frac{1}{\ln x} \right)$ corresponds to the case $\infty - \infty$. On transforming it to the form

$$A = \lim_{x \rightarrow 1} \frac{x \ln x - x + 1}{(x-1) \ln x}$$

we arrive at the case $\frac{0}{0}$. Now, using L'Hôpital's rule we finally obtain

$$A = \lim_{x \rightarrow 1} \frac{\ln x}{\frac{x-1}{x} + \ln x} = \lim_{x \rightarrow 1} \frac{\frac{1}{x}}{\frac{1}{x^2} + \frac{1}{x}} = \frac{1}{2}$$

(3) Consider the limit $A = \lim_{x \rightarrow 0} (\cos x)^{\frac{1}{x^2}}$ which belongs to the case 1^∞ . Taking the logarithm we reduce it to the limit $A_1 = \lim_{x \rightarrow 0} \ln (\cos x)^{\frac{1}{x^2}}$ which can be rewritten as $A_1 = \lim_{x \rightarrow 0} \frac{\ln \cos x}{x^2}$; we thus arrive at the case $\frac{0}{0}$. L'Hôpital's rule yields

$$A_1 = \lim_{x \rightarrow 0} \frac{-\tan x}{2x} = -\frac{1}{2}$$

whence, since $A_1 = \ln \lim_{x \rightarrow 0} (\cos x)^{\frac{1}{x^2}} = \ln A$, we find $A = e^{A_1} = e^{-\frac{1}{2}} = \frac{1}{\sqrt{e}}$.

(4) The limit $A = \lim_{x \rightarrow 0} (\cot x)^{\frac{1}{\ln x}}$ belongs to the type ∞^0 and its computation is reducible to the limit $A_1 = \lim_{x \rightarrow 0} \ln (\cot x)^{\frac{1}{\ln x}}$. The latter can be written as $A_1 = \lim_{x \rightarrow 0} \frac{\ln \cot x}{\ln x}$ (the case $\frac{\infty}{\infty}$). On applying L'Hôpital's rule we find

$$A_1 = \lim_{x \rightarrow 0} \frac{\frac{1}{\cot x \sin^2 x}}{\frac{1}{x}} = \lim_{x \rightarrow 0} \frac{-x}{\cos x \sin x} = -1$$

whence $A = e^{A_1} = e^{-1} = \frac{1}{e}$.

(5) Now let us determine $A = \lim_{x \rightarrow 0} x^x$ (the case 0^0). Here we take the limit $A_1 = \lim_{x \rightarrow 0} \ln x^x$. On transforming it we obtain

$$A_1 = \lim_{x \rightarrow 0} x \ln x = 0$$

(see Example 1 for $n=1$) and thus find $A = e^{A_1} = e^0 = 1$.

64. Asymptotes of Curves. When studying the behaviour of a function as its argument tends to plus or minus infinity we deal with the *branches* of its graph *receding to infinity*. Such branches are also encountered when we investigate a function in the vicinity of its infinite discontinuity (see Sec. 33). The determination of infinite branches of the graph of a function is needed for a correct and more complete representation of the graph as a whole, that is, for describing the behaviour of the function throughout its domain of definition.

We shall approach this problem from a geometrical point of view and introduce, in this connection, the general definition of an *asymptote* of a curve (some particular types of asymptotes were dealt with in foregoing sections).

Definition. A straight line Γ is called an *asymptote* of a curve L if the distance between the moving point of the curve L and the line Γ tends to zero as the distance from this point to the origin increases indefinitely.

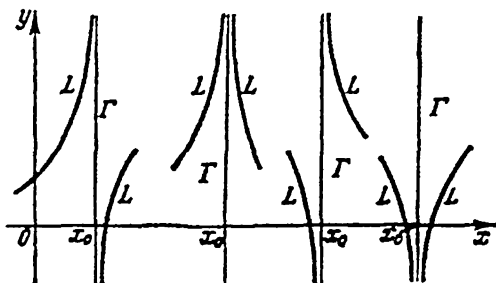


Fig. 83

One should distinguish between *vertical* and *inclined* asymptotes.

(1) Let a curve $y=f(x)$ have a *vertical* asymptote. The equation of such an asymptote is of the form $x=x_0$, and hence, according to the definition of an asymptote, there must be $f(x) \rightarrow \infty$ for $x \rightarrow x_0$. Conversely, if x_0 is a point of infinite (at least, one-sided) discontinuity of the function $y=f(x)$ the straight line $x=x_0$ serves as a vertical asymptote of the graph of the function $y=f(x)$.

Thus, if $\lim_{x \rightarrow x_0} f(x) = \infty$ the curve $y=f(x)$ has the vertical asymptote $x=x_0$.

The concrete disposition of a branch of a curve receding to infinity relative to its vertical asymptote $x=x_0$ is studied by finding out whether $f(x)$ becomes positively or negatively infinite as x approaches x_0 , remaining less than x_0 , that is from the left, or remaining greater than x_0 , that is from the right. What has been said is demonstrated in Fig. 83 where the possible cases are

shown. It may occur that the graph only approaches its asymptote from one side. For instance, this is the case for the function $y = \ln x$.

(2) Let a curve L having an equation $y = f(x)$ possess an inclined asymptote Γ with an equation of the form $y = ax + b$ ($a \neq 0$). By the definition of an asymptote, the distance MN_1 from the point M of the curve L to the asymptote Γ (see Fig. 84) tends to zero as $x \rightarrow \infty$. Instead of the distance MN_1 , it is more convenient to consider the line segment MN , that is the difference between the ordinate of the point M and that of the point N lying on the straight line Γ and having the same abscissa as the point M . We have (see

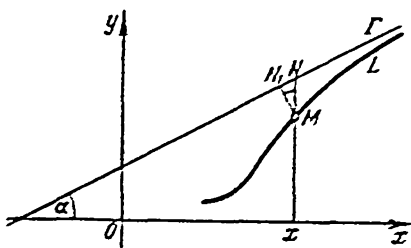


Fig. 84

Fig. 84) $MN = \frac{MN_1}{\cos \alpha}$, i.e. $MN_1 = MN \cos \alpha$, where α is the angle between the asymptote and the x -axis, and therefore the distances MN_1 and MN tend to zero simultaneously.

The ordinate of the point M being equal to the value of the function $f(x)$ and that of the point N to the value of the linear function $ax + b$, we have

$$MN = |f(x) - (ax + b)|$$

This means that if a curve $y = f(x)$ has an asymptote $y = ax + b$ then

$$\lim_{x \rightarrow \infty} [f(x) - (ax + b)] = 0$$

and, conversely, if the difference $f(x) - (ax + b)$ tends to zero, as $x \rightarrow \infty$, the straight line $y = ax + b$ is an asymptote of the curve $y = f(x)$.

Thus, the question of the existence and determination of an inclined asymptote of a curve $y = f(x)$ reduces to the question of the existence and determination of two numbers a and b such that

$$\lim_{x \rightarrow \infty} [f(x) - ax - b] = 0 \quad (*)$$

If such numbers exist we have

$$f(x) = ax + b + \alpha(x)$$

where $\alpha(x)$ is an infinitesimal as $x \rightarrow \infty$. On dividing both sides of this relation by x and passing to the limit as $x \rightarrow \infty$ we obtain

$$\lim_{x \rightarrow \infty} \frac{f(x)}{x} = \lim_{x \rightarrow \infty} \left(a + \frac{b}{x} + \frac{\alpha(x)}{x} \right)$$

Now, since $\frac{b}{x} \rightarrow 0$ and $\frac{\alpha(x)}{x} \rightarrow 0$, the number a is specified by the relation

$$a = \lim_{x \rightarrow \infty} \frac{f(x)}{x} \quad (**)$$

After a has been determined the number b is found by computing the limit

$$\lim_{x \rightarrow \infty} [f(x) - ax] = b \quad (***)$$

It is apparent that, conversely, if the limits (**) and (***) exist and the numbers a and b are found the basic condition for the existence of an inclined asymptote is fulfilled, and thus,

If the ratio $\frac{f(x)}{x}$ tends to a finite limit a as $x \rightarrow \infty$ and if the difference $f(x) - ax$ has a finite limit b as $x \rightarrow \infty$ the curve $y = f(x)$ possesses the inclined asymptote $y = ax + b$.

In particular, if the function $f(x)$ tends to a finite limit as $x \rightarrow \infty$, that is

$$\lim_{x \rightarrow \infty} f(x) = b$$

then, obviously, $a = 0$, and the curve $y = f(x)$ has a *horizontal* asymptote (regarded as a special case of the inclined asymptote) parallel to Ox , namely, $y = b$. If both a and b are equal to zero the x -axis itself serves as a horizontal asymptote.

If at least one of these limits does not exist the curve $y = f(x)$ has no inclined (including horizontal) asymptotes.

For example, taking the curve $y = x + \ln x$ we find

$$\lim_{x \rightarrow \infty} \frac{x + \ln x}{x} = \lim_{x \rightarrow \infty} \left(1 + \frac{\ln x}{x}\right) = 1, \text{ i.e. } a = 1$$

At the same time,

$$\lim_{x \rightarrow \infty} (x + \ln x - 1 \cdot x) = \infty$$

and hence there is no inclined asymptote in this case.

Note that in the latter example we have the asymptotic relation $x + \ln x \sim x$ for $x \rightarrow \infty$ (see Sec. 37). Generally, it follows from formula (**) that if $a \neq 0$ then $f(x) \sim ax$ for $x \rightarrow \infty$. As we see, such an asymptotic relation by far not always guarantees the existence of an asymptote $y = ax + b$.

The *asymptotic behaviour* of a function may be of different character when x becomes *positively* or *negatively* infinite, and therefore the cases $x \rightarrow +\infty$ and $x \rightarrow -\infty$ should be treated separately. It may happen that a graph possesses an asymptote as $x \rightarrow +\infty$ and another asymptote as $x \rightarrow -\infty$ or only one of them. In a particular case the limits (**) and (***) obtained,

respectively, for $x \rightarrow +\infty$ and for $x \rightarrow -\infty$ may coincide, and then both asymptotes merge to form a common straight line.

On finding an asymptote $y = ax + b$ we can investigate the sign of the expression $\alpha(x) = f(x) - ax - b$ for large values of $|x|$ and thus determine the mutual disposition of the branch of the curve receding to infinity and its asymptote. If x becomes sufficiently large in its absolute value, then, depending on the sign of $\alpha(x) = f(x) - ax - b$, the branch in question lies entirely above its asymptote (if $\alpha > 0$) or below it (if $\alpha < 0$) or intersects it infinitely many times (if the sign of α alternates indefinitely).

Examples. (1) Let $y = \frac{x^3}{x^2 - 1}$. The graph of this function has two vertical asymptotes: $x = -1$ and $x = 1$. To determine its inclined asymptote we compute the limits

$$a = \lim_{x \rightarrow \infty} \frac{x^3}{x(x^2 - 1)} = 1 \text{ and } b = \lim_{x \rightarrow \infty} \left[\frac{x^3}{x^2 - 1} - x \right] = \lim_{x \rightarrow \infty} \frac{x}{x^2 - 1} = 0$$

Thus, the bisector of the first quadrant $y = x$ is the only inclined asymptote the graph possesses.

Generally, if a function in question is a rational fraction the limits a and b can be looked for as x tends to infinity in an arbitrary manner (see Example 5 in Sec. 29).

(2) Let $y = x^n e^{-x}$ ($n > 0$). The graph has no vertical asymptote. Example 9 in Sec. 63, I shows that $\lim_{x \rightarrow +\infty} x^n e^{-x} = 0$, and, consequently, $a = 0$ and $b = 0$. Thus, the graph has the x -axis as its asymptote for $x \rightarrow +\infty$. As $x \rightarrow -\infty$, the graph has no asymptote since $\lim_{x \rightarrow -\infty} \frac{x^n e^{-x}}{x} = \infty$. On the contrary, the function $y = x^n e^x$ ($n > 0$) has an asymptote as $x \rightarrow -\infty$ (which is again the x -axis) and has no asymptote as $x \rightarrow +\infty$.

We can now complete the investigation of the function $y = \sqrt[3]{x^2} e^x$ (treated in Secs. 60, III and 62, IV) for its entire domain of definition, i.e., throughout Ox , by determining its asymptotes for $x \rightarrow +\infty$ and $x \rightarrow -\infty$. The above example shows that the x -axis is an asymptote of the graph as $x \rightarrow -\infty$ whereas for $x \rightarrow +\infty$ there is no asymptote. Therefore, the graph is of the shape depicted in Fig. 82.

(3) The curve $y = \frac{\sin x}{x}$ has the asymptote $y = 0$ which it cuts infinitely many times. Let the reader construct the graph of this function.

(4) Let us determine the asymptotes of the graph of the function $y = x \arctan x$. Since $\frac{y}{x} = \arctan x$ tends to $\frac{\pi}{2}$ as $x \rightarrow +\infty$ and

to $-\frac{\pi}{2}$ as $x \rightarrow -\infty$ we should compute the limits $\lim_{x \rightarrow +\infty} \left(y - \frac{\pi}{2}x\right)$ and $\lim_{x \rightarrow -\infty} \left(y + \frac{\pi}{2}x\right)$.

The application of L'Hospital's rule results in

$$\lim_{x \rightarrow +\infty} \left(y - \frac{\pi}{2}x\right) = \lim_{x \rightarrow +\infty} \frac{\arctan x - \frac{\pi}{2}}{\frac{1}{x}} = \lim_{x \rightarrow +\infty} \frac{\frac{1}{1+x^2}}{-\frac{1}{x^2}} = -1$$

and, similarly, $\lim_{x \rightarrow -\infty} \left(y + \frac{\pi}{2}x\right) = -1$. Hence, there are two asymptotes: $y = \frac{\pi}{2}x - 1$ which is approached as $x \rightarrow +\infty$ and $y = -\frac{\pi}{2}x - 1$ which is approached as $x \rightarrow -\infty$. To represent the whole graph more correctly we shall consider the first and second derivatives of the given function,

The first derivative is

$$y' = \arctan x + \frac{x}{1+x^2}$$

Both terms on the right are positive for $x > 0$, and, consequently, $y' > 0$, and the function increases on the interval $(0, \infty)$. For $x < 0$ these terms are both negative, i.e. $y' < 0$, and the function decreases on the interval $(-\infty, 0)$ (the latter result is also implied by the fact that the function is even). Thus, $x=0$ is a point of minimum, the value of the function at this point being $y_{x=0} = 0$.

The second derivative

$$y'' = \frac{1}{1+x^2} + \frac{1-x^2}{(1+x^2)^2} = \frac{2}{(1+x^2)^2}$$

is everywhere positive, which means that the graph is concave at all its points. These considerations lead to the schematic graph of the function $y = x \arctan x$ shown in Fig. 85.¹

For sufficiently large values of $|x|$ the function $x \arctan x$ can be approximately replaced by linear functions (as we say, can be *linearized*):

$$x \arctan x \approx \frac{\pi}{2}x - 1 \quad \text{for } x > 0$$

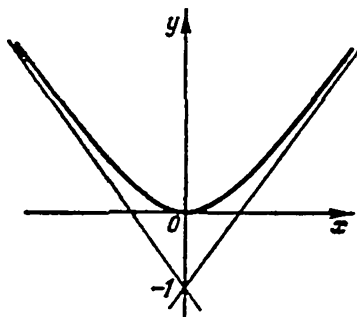


Fig. 85

¹ We leave it to the reader to check that the graph does not intersect its asymptotes.

and

$$x \arctan x \approx -\frac{\pi}{2}x - 1 \quad \text{for } x < 0$$

65. General Scheme for the Investigation of the Graph of a Function. Let us now elaborate a general scheme convenient for studying functions and curves.

For a given function $y=f(x)$ this scheme reduces to four stages at which we determine, in succession, the following elements of the behaviour of the function:

- I. (1) The domain of definition of the function.
- (2) Its points of discontinuity and intervals of continuity.
- (3) The behaviour of the function in the vicinity of the points of discontinuity and its vertical asymptotes.
- (4) The points at which the graph cuts the axes of coordinates.
- (5) The character of symmetry of the graph, that is whether the function is even or odd (or neither).
- (6) The periodicity of the graph.

II. The intervals of monotonicity of the function and the points of extremum together with the extremal values of the function.

III. The intervals of convexity and concavity and the points of inflection.

IV. The behaviour of the function at infinity and its inclined (and, in particular, horizontal) asymptotes.

At the first stage we give a general description of the peculiarities of the function and of its graph. In this investigation one should bear in mind the essence of the problem leading to the function under consideration. For instance, it may turn out that it is allowable to take a narrower interval of variation of the independent variable than the domain of definition of the function. It is sometimes convenient to realize the fourth stage simultaneously with the first one at which the general character of the behaviour of the function is studied.

It sometimes suffices to give a verbal description of the characteristic elements of the behaviour of the function, but a more precise specification taking into account all the peculiarities is obtained more economically and visually if the functional relationship is represented graphically. Therefore, *in the realization of these stages it is advisable to accompany it by a gradual construction of the graph of the function.* First of all, one should mark on the axis of abscissas the characteristic points which are *the points of discontinuity, the zeros, the points of extremum and the abscissas of the points of inflection.* The points of the graph corresponding to these values of the argument are plotted in the xy -plane.

In the intervals between these characteristic points the behaviour of the function does not change sharply and it varies gradually, but to specify the course of the variation more precisely it is advisable to take some ordinary points within these intervals, to compute the corresponding values of the function and to plot the new points of the graph. The greater the number of such intermediate points, the more complete the representation of the graph given by the line drawn through all the constructed points in the xy -plane.

Examples. (1) Let us investigate the function $y = \frac{x^2}{1+x}$.

I. The function is defined and continuous throughout Ox except the point $x = -1$ at which it has an infinite discontinuity. Consequently, the straight line $x = -1$ is a vertical asymptote; furthermore,

$$\lim_{\substack{x \rightarrow -1 \\ x < -1}} y = -\infty \quad \text{and} \quad \lim_{\substack{x \rightarrow -1 \\ x > -1}} y = +\infty$$

The only zero of the function is the point $x = 0$.

II. The first derivative is

$$y' = \frac{x^2 + 2x}{(1+x)^2} = \frac{x(x+2)}{(1+x)^2}$$

It turns into zero for $x = 0$ and for $x = -2$; in the interval $(-\infty, -2)$ it is positive and in the intervals $(-2, -1)$ and $(-1, 0)$ negative (the function is not defined at the point $x = -1$ itself) while in the interval $(0, \infty)$ it is again positive. Consequently, the function increases in the first interval, decreases in the second and third intervals and increases in the fourth interval; $x = -2$ is a point of maximum, the maximum value of the function being -4 ; the point $x = 0$ is a point of minimum, and the minimum value of the function is equal to zero.

III. Now we find the second derivative:

$$y'' = \frac{2}{(1+x)^3}$$

It does not turn into zero anywhere but changes sign from $(-)$ to $(+)$ as x passes through the point $x = -1$. Thus, the second derivative is negative in the interval $(-\infty, -1)$ and positive in the interval $(-1, \infty)$. In the former interval the graph of the function is convex and in the latter concave.

IV. Since

$$\lim_{x \rightarrow \infty} \frac{y}{x} = \lim_{x \rightarrow \infty} \frac{x}{1+x} = 1$$

and

$$\lim_{x \rightarrow \infty} (y - 1 \cdot x) = \lim_{x \rightarrow \infty} \left(-\frac{x}{1+x} \right) = -1$$

there exists an inclined asymptote: $y = x - 1$. The difference

$$\alpha(x) = \frac{x^2}{1+x} - (x-1) = \frac{1}{1+x}$$

being positive for $x > -1$ and negative for $x < -1$, the graph lies above the asymptote $y = x - 1$ on the right of the straight line $x = -1$ and below the asymptote on the left of the line $x = -1$.

If we put, approximately,

$$y = \frac{x^2}{1+x} \approx x - 1$$

the absolute error is $\alpha = \left| \frac{1}{1+x} \right|$ and the relative error is $\delta = \frac{1}{x^2}$. For $|x| > 10$ the relative error does not exceed 1%. Thus, for $|x| > 10$ the function $y = \frac{x^2}{1+x}$ can be practically replaced by the

linear function $y = x - 1$, which simplifies the problem.

Finally, taking into account the results obtained we readily construct the graph (see Fig. 86) providing a correct description of the character of variation of the function throughout Ox . The application of the methods of analytic geometry shows that the curve

$y = \frac{x^2}{1+x}$ is a hyperbola.

(2) Let us study the function $y = e^{-\frac{x^2}{2}}$.

I. The domain of definition and the interval of continuity coincide with Ox .

Since $y > 0$ for all the values of x , the entire graph lies above the axis of abscissas. The function is even, and

its graph is symmetric about Oy .

As $x \rightarrow \infty$ the function tends to zero, and therefore the axis of abscissas is a horizontal asymptote.

II. The first derivative is $y' = -xe^{-\frac{x^2}{2}}$. In the interval $(-\infty, 0)$ it is greater than zero, and the function increases, while in the interval $(0, \infty)$ the derivative is less than zero, and the function decreases. At the point $x=0$ there is a maximum, the maximum value $y_{x=0} = 1$ being the greatest value of the function.

The information obtained is sufficient for constructing schematically the graph of the function, which immediately shows that there are points of inflection. We are going to find them now.

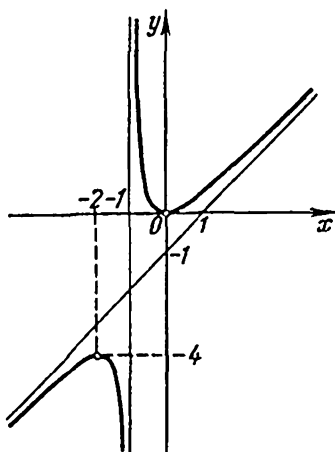


Fig. 86

III. Let us compute the second derivative: $y'' = e^{-\frac{x^2}{2}}(x^2 - 1)$. The roots of the equation $y'' = 0$ are the points $x = \pm 1$. The sign of the second derivative is completely specified by that of the second factor entering into its expression: the derivative is positive in the intervals $(-\infty, -1)$ and $(1, \infty)$ and negative in the interval $(-1, 1)$. Hence, in the first two intervals the graph of

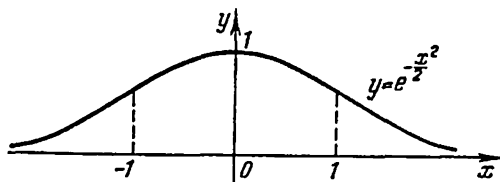


Fig. 87

the function is concave and in the third interval (lying between them) it is convex, $x = \pm 1$ being the points of inflection. The value of the function at the points of inflection coincide and are equal to $y_{x=\pm 1} = e^{-0.5} \approx 0.6065$.

The final sketch of the graph is given in Fig. 87. The function considered in this example plays an important role in the *theory of probability*.

§ 4. Curvature

66. Differential of Arc Length. Let us begin with the well-known notion of the circumference of a circle; by definition, it is equal to the limit of the perimeter of the regular inscribed polygon of the circle as the number of its sides increases indefinitely. The arc length of any other curve is defined analogously:

The arc length is the limit (provided it exists) to which the length of the polygonal line inscribed in the arc tends as its segments are refined indefinitely.

In the integral calculus we shall show that if a given function $y = f(x)$ and its derivative $f'(x)$ are continuous in a closed interval $[a, b]$ the

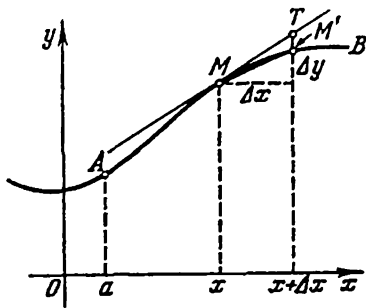


Fig. 88

above limit exists for the arc $\overline{AB}$ (see Fig. 88) corresponding to that interval, i.e. this arc has length. From the definition it of course follows that any arc $\overline{AM}$ where $M \in \overline{AB}$ also has length. The length of the arc $\overline{AM}$ depends on the location of the point M

and is a function of the abscissa x of this point which we shall denote by $s(x)$:

$$\widetilde{AM} = s(x)$$

The problem of practical computation of the arc length, that is of the function $s(x)$, corresponding to a given equation $y=f(x)$ will be considered in Sec. 102. Here we shall confine ourselves to finding the differential of this function referred to as the *differential (or element) of arc length*. Since

$$ds = s'(x) dx$$

it is convenient to begin with the determination of the derivative of the function $s(x)$, that is of the limit of the ratio of the length of the arc $\widetilde{MM'}$, corresponding to the interval $[x, x + \Delta x]$, to Δx :

$$s'(x) = \frac{ds}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\widetilde{MM'}}{\Delta x} \quad (*)$$

It is assumed that both the function $f(x)$ and its first derivative $f'(x)$ are continuous, which guarantees the absence of corner points and cusps on the arc $\widetilde{AB}$ of the line $y=f(x)$. A curve of that kind is said to be *smooth*. Now we assume without proof the following geometrical fact: *the limit of the ratio of the arc length of a smooth curve to the length of its chord as the arc length tends to zero is equal to unity*: $\lim_{\Delta x \rightarrow 0} \frac{\widetilde{MM'}}{\overline{MM'}} = 1$.

For a circle this property is expressed by the limiting relation

$$\lim_{\alpha \rightarrow 0} \frac{\sin \alpha}{\alpha} = 1$$

studied in Sec. 30.

On the basis of the properties of equivalent infinitesimals (see Sec. 35, II), the arc $\widetilde{MM'}$ in the limit (*) can be replaced by the chord $\overline{MM'} = \sqrt{\Delta x^2 + \Delta y^2}$ equivalent to it as infinitesimal. Therefore,

$$\frac{ds}{dx} = \lim_{\Delta x \rightarrow 0} \frac{\sqrt{\Delta x^2 + \Delta y^2}}{\Delta x} = \lim_{\Delta x \rightarrow 0} \sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} = \sqrt{1 + y'^2}$$

whence

$$ds = \sqrt{1 + y'^2} dx.$$

On transforming this expression we arrive at the following simple formula for the differential of arc length which can be easily memorized:

$$ds = \sqrt{dx^2 + dy^2}$$

It now follows that

the differential of arc length is equal to the length of the corresponding segment of the tangent line at the initial point of the arc (see the segment MT in Fig. 88).

Observe, that, although the length of the chord $\overline{MM'}$ and that of the segment of the tangent MT are both equivalent to the infinitesimal arc length $\overline{MM'}$, the segment MT is proportional to dx whereas the segment MM' is not, and therefore the differential of arc length is the length of the segment of the tangent to the arc but not the length of the chord subtending that arc.

67. Curvature.

1. The first derivative $f'(x)$ of a function $f(x)$ is a simple characteristic of the curve $y=f(x)$ specifying its *direction* (more precisely, the direction of its tangent). It turns out that the second derivative $f''(x)$ is closely related to another quantitative characteristic of the curve known as its *curvature* which characterizes the degree of the "bending" of the curve.

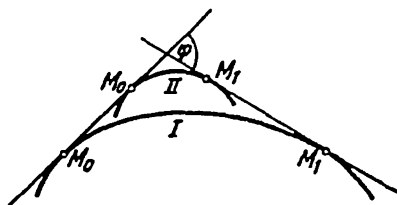


Fig. 89

In what follows we always assume that the functions we deal with and its derivatives $f'(x)$ and $f''(x)$ are continuous.

Consider an arc $\widetilde{M_0M_1}$ (Fig. 89). The angle φ (measured in radians) between the tangents drawn at the end points M_0 and M_1 of the arc $\widetilde{M_0M_1}$ is termed the *angle of contingence* of that arc.

The angle of contingence φ characterizes, to a certain extent, how the arc $\widetilde{M_0M_1}$ bends since it is the angle through which the tangent turns as the point of tangency moves from the initial point M_0 of the arc to the terminal point M_1 ; the greater the angle of contingence, the greater the bending of the arc. At the same time, an angle of contingence of a given magnitude may characterize two arcs whose degrees of bending are essentially different. For instance, in Fig. 89, I the angle of contingence φ corresponds to the greater arc (whose bending is less) while in Fig. 89, II it corresponds to the smaller arc (whose bending is greater). Therefore, in order to characterize properly the degree of the bending of a curve it is natural to relate the angle of contingence to unit arc length.

Definition. The ratio of the angle of contingence of an arc to its length is called the *average curvature* of that arc:

$$K_{av} = \frac{\varphi}{\widetilde{M_0M_1}}$$

On the other hand, for different parts of one and the same arc the average curvature may be different, and to avoid the ambi-

guity and obtain a more precise characteristic of the bending of the arc we introduce the notion of the *curvature at a given point* M_0 proceeding from the fact that, the smaller the arc M_0M_1 , the more precise the quantitative characteristic of the bending of the arc in the vicinity of the point M_0 given by the average curvature:

Definition. The *curvature* K of a curve at its point M_0 is the limit (provided it exists) of the average curvature K_{av} of the arc M_0M_1 as the terminal point of the arc M_1 tends to its initial point M_0 :

$$K = \lim_{M_1 \rightarrow M_0} K_{av} = \lim_{M_1 \rightarrow M_0} \frac{\varphi}{\overline{M_0M_1}}$$

In particular, for a circle of radius r we obtain (see Fig. 90):

$$K_{av} = \frac{\varphi}{\overline{M_0M_1}} = \frac{\varphi}{r\varphi} = \frac{1}{r}$$

Thus, the average curvature of a circle is constant, and hence its curvature is also constant at each point and is the reciprocal of its radius. This is completely coherent with our intuitive knowledge of the character of the bending of a circle.

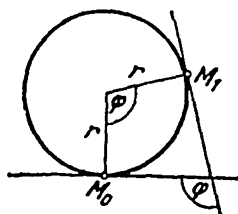


Fig. 90

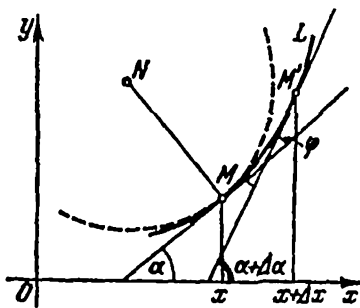


Fig. 91

Let us proceed to find the expression for the curvature of a line specified by an equation $y = f(x)$ in Cartesian coordinates (the corresponding construction is shown in Fig. 91). On the given curve we take the points M and M' whose abscissas are, respectively, x and $x + \Delta x$, and denote by α and $\alpha + \Delta\alpha$ the angles of inclination of the tangents at the points M and M' to the axis of abscissas, the arc length of $\overline{MM'}$ being Δs . Since the angle of contingence φ of the arc $\overline{MM'}$ is equal to $|\Delta\alpha|$,¹ we have

$$K_{av} = \left| \frac{\Delta\alpha}{\Delta s} \right| = \left| \frac{\frac{\Delta\alpha}{\Delta x}}{\frac{\Delta s}{\Delta x}} \right|$$

¹ We have $\Delta\alpha > 0$ for a concave arc and $\Delta\alpha < 0$ for a convex arc. That is why we write $\varphi = |\Delta\alpha|$.

The condition $M' \rightarrow M$ is equivalent to $\Delta x \rightarrow 0$, and therefore

$$K = \lim_{\Delta x \rightarrow 0} K_{\sigma\sigma} = \left| \frac{\frac{d\alpha}{dx}}{\frac{ds}{dx}} \right| = \left| \frac{\frac{d\alpha}{dx}}{\sqrt{1+y'^2}} \right|$$

(see Sec. 66). From the equality $\tan \alpha = y'$ we obtain $\alpha = \arctan y'$, whence

$$\frac{d\alpha}{dx} = \frac{y''}{1+y'^2}$$

and, consequently,

$$K = \left| \frac{y''}{(1+y'^2)\sqrt{1+y'^2}} \right| = \left| \frac{y''}{(1+y'^2)^{3/2}} \right| \quad (*)$$

Formula (*) specifies the curvature as a function of the abscissa of the moving point of the curve. If the curve is represented by parametric equations $x = \varphi(t)$, $y = f(t)$ we have (see Sec. 54, III)

$$y' = \frac{f'(t)}{\varphi'(t)} \quad \text{and} \quad y'' = \frac{f''(t)\varphi'(t) - f'(t)\varphi''(t)}{[\varphi'(t)]^3}$$

whence

$$K = \left| \frac{f''(t)\varphi'(t) - f'(t)\varphi''(t)}{[\varphi'^3(t) + f'^2(t)]^{3/2}} \right|$$

or, in the contracted notation,

$$K = \left| \frac{y'x'' - y''x'}{(x'^2 + y'^2)^{3/2}} \right| \quad (**)$$

Formula (**) expresses the curvature as a function of the value of the parameter t corresponding to the variable point of the curve. If $t = x$ formula (**) turns into (*).

The curvatures of a circle and of a straight line are constant; as follows from formula (*), the curvature of a straight line is equal to zero, which is coherent with the intuitive idea that the straight line does not bend. In the general case, the curvature of a line varies from point to point.

Example. Let us determine the curvature of the parabola $y = x^2$ at its variable point. We have $y' = 2x$ and $y'' = 2$, which yields $K = \frac{2}{(\sqrt{1+4x^2})^3}$. In particular, the curvature of the parabola at its vertex is equal to 2.

II. Radius, centre and circle of curvature. Evolute and involute. Let us draw the normal to the line L (shown in Fig. 91) at the

point M (in the direction of its concavity) and lay off from the point M the line segment $MN = \rho$ along this normal where ρ is the reciprocal of the curvature K :

$$\rho = \frac{1}{K}$$

Definition. The line segment MN is termed the *radius of curvature*, the point N is the *centre of curvature* and the circle of radius MN with centre at the point N is called the *circle of curvature* of the line L at its point M .

The curvature of the circle of curvature obviously coincides with that of the given line at the point M .

For the radius of curvature we obtain the formula

$$\rho = \left| \frac{(1+y'^2)^{3/2}}{y''} \right|$$

The radius of curvature like the curvature itself serves as the measure of the bending of the given curve: the smaller the radius of curvature at the given point, the greater the bending of the curve at that point. For a straight line we assume, conditionally, that at every point its radius of curvature is infinite; the same refers to any point of a curve at which the curvature turns into zero. Later on we shall elucidate the role of the radius of curvature of a line in the investigation of a curvilinear motion in mechanics (see Sec. 71).

It can be shown that if an infinitesimal arc of a curve is replaced, in the vicinity of a given point, by the corresponding arc of its circle of curvature there appears an error of a higher order of smallness than that in the replacement of the arc by the corresponding segment of its tangent line. Thus, the arc of the circle of curvature is a better approximation to a small arc of a curve than the segment of its tangent.

With each point M of a curve L there is associated the centre of curvature N corresponding to the point M .

Definition. The locus of the centres of curvature L' of a given curve L is called its *evolute*, the curve L itself being called the *evolvent* of its evolute L' .

Below we present without proof some techniques for an approximate construction of the evolute of a given evolvent and of the evolvent of a given evolute.

(1) Every normal to the evolvent is a tangent to the evolute at the corresponding centre of curvature; this means that the evolute *envelopes* the whole family of normals to the evolvent, and hence, to construct, approximately, the evolvent it is suffi-

cient to draw a large number of normals to the evolute L and to draw the envelope of these lines which thus is the evolute L' (see Fig. 92).

(2) If a flexible unstretchable taut thread is wound off from the contour of a given convex line L' (serving as an evolute) the end of the thread (and, generally, its every fixed point) describes an evolute (see the line L in Fig. 92). (The evolute is also termed the *involute*). Winding off a taut thread is equivalent to rolling a straight line (without sliding) upon the given curve L' : every fixed point

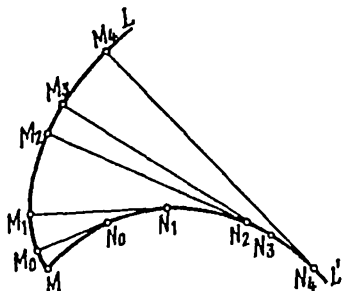


Fig 92

of this straight line describes an evolute L of the curve L' . It follows that a given evolute L' possesses an infinitude of evolutes L . At the same time a given curve regarded as an evolute has a single evolute.

§ 5. Space Curves. Vector Function of a Scalar Argument¹

68. Space Curves. As was shown in Sec. 46, a curve lying in the xy -plane can be specified by parametric equations $x=x(t)$, $y=y(t)$. Analogously, a space curve can be represented parametrically by means of equations

$$x=x(t), \quad y=y(t), \quad z=z(t) \quad (*)$$

As the parameter t varies, the moving point $M(x, y, z)$ whose coordinates are specified by formulas $(*)$ describes a curve.

If all the three equations $(*)$ are linear, that is,

$$x=mt+a, \quad y=nt+b, \quad z=pt+c$$

we obtain *parametric equations of a straight line* well known from analytic geometry. On eliminating the parameter t we arrive at *canonical equations of the straight line*:

$$\frac{x-a}{m} = \frac{y-b}{n} = \frac{z-c}{p}.$$

The tangent line to a space curve at its point M is defined as in the case of a plane curve, that is, as the *limiting position of the secant passing through the point M and another point M' of that curve placed closely to the former on condition that the point M' tends to the given point M* (see Sec. 40).

¹ In § 5 it is supposed that the reader is familiar with fundamentals of linear algebra and solid analytic geometry.

Let us derive canonical equations of the tangent to a curve, specified by parametric equation (*), at its point $M_0(x_0, y_0, z_0)$ corresponding to a value $t=t_0$ of the parameter. We shall assume that each of the functions (*) possesses a derivative for $t=t_0$ and that these derivatives do not vanish simultaneously.

The equations of the secant passing through the points

$$M_0(x_0, y_0, z_0) \text{ and } M'(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$$

can be written in the form

$$\frac{x-x_0}{\Delta x} = \frac{y-y_0}{\Delta y} = \frac{z-z_0}{\Delta z}$$

On dividing all the denominators by the corresponding increment Δt of the parameter and passing to the limit as $\Delta t \rightarrow 0$ we receive the desired equations

$$\frac{x-x_0}{x'(t_0)} = \frac{y-y_0}{y'(t_0)} = \frac{z-z_0}{z'(t_0)} \quad (**)$$

where $x_0 = x(t_0)$, $y_0 = y(t_0)$, and $z_0 = z(t_0)$. As has already been mentioned, it is supposed that at least one of the derivatives $x'(t_0)$, $y'(t_0)$ and $z'(t_0)$ is different from zero. If this condition is violated the point (x_0, y_0, z_0) is called a *singular point* of the curve; we shall not treat such points.

A straight line perpendicular to the tangent and passing through the point of tangency $M_0(x_0, y_0, z_0)$ is called, as in the case of a plane curve, a *normal* to the given curve at that point. It is obvious that, in contrast to a plane curve (whose tangents and normals are lying in the given plane), a space curve has an infinity of normals at its every point; all the normals at a given point lie in a common plane passing through that point of tangency perpendicularly to the tangent line.

Definition. The plane perpendicular to the tangent line to a given curve at a given point of tangency is called the *normal plane* of the curve at that point.

According to the well-known perpendicularity condition of analytic geometry, the equation of the normal plane as plane perpendicular to the straight line (**) and passing through the point $M_0(x_0, y_0, z_0)$ is written as

$$x'(t_0)(x-x_0) + y'(t_0)(y-y_0) + z'(t_0)(z-z_0) = 0$$

Example. Let a curve be specified by the equations

$$x = t^2 - 1, \quad y = t^2 + 2, \quad z = 3t - 1$$

It is required to form equations of the tangent line and of the normal plane of this curve at the point $M_0(0, 3, 2)$ corresponding to the value $t=1$ of the parameter. We have $x'_{t=1}=2$, $y'_{t=1}=2$

and $z'_{t=1}=3$, and hence, the equations of the tangent are

$$\frac{x}{2} = \frac{y-3}{2} = \frac{z-2}{3}$$

while the equation of the normal plane is

$$2x + 2y + 3z - 12 = 0$$

For our further aims we need the differentiation formula for the arc length of a space curve specified by equations (*). It is completely analogous to the corresponding formula for a plane curve (see Sec. 66). As before, we shall deal with *smooth* curves for which the functions (*) possess continuous derivatives not vanishing simultaneously. For such a curve the ratio of the length Δs of the arc contained between the points $M(x, y, z)$ and $M'(x+\Delta x, y+\Delta y, z+\Delta z)$ to the length of the chord joining these points tends to unity as the point M' approaches the point M . The length of the chord being $\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$, we can write

$$\frac{\Delta s}{\sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}} \rightarrow 1$$

On dividing the numerator and the denominator by the increment Δt of the parameter corresponding to the increments $\Delta x, \Delta y, \Delta z$ of the coordinates we receive

$$\frac{\frac{\Delta s}{\Delta t}}{\sqrt{\left(\frac{\Delta x}{\Delta t}\right)^2 + \left(\frac{\Delta y}{\Delta t}\right)^2 + \left(\frac{\Delta z}{\Delta t}\right)^2}} \rightarrow 1$$

Now, passing to the limit as $\Delta t \rightarrow 0$, we obtain

$$\frac{\frac{ds}{dt}}{\sqrt{x'^2 + y'^2 + z'^2}} = 1$$

that is,

$$ds = \sqrt{x'^2 + y'^2 + z'^2} dt \quad \text{or} \quad ds = \sqrt{dx^2 + dy^2 + dz^2}$$

As was shown above the direction vector $\mathbf{T}$ of the tangent line to the curve at the point M has the projections $\{x', y', z'\}$, where, for the sake of brevity, x' designates $x'(t)$ etc. The direction cosines of the vector $\mathbf{T}$ are

$$\cos \alpha = \frac{x'}{\sqrt{x'^2 + y'^2 + z'^2}}, \quad \cos \beta = \frac{y'}{\sqrt{x'^2 + y'^2 + z'^2}},$$

$$\cos \gamma = \frac{z'}{\sqrt{x'^2 + y'^2 + z'^2}}$$

Using the formula for the differential of arc length we can rewrite the formulas for the direction cosines in the form

$$\cos \alpha = \frac{dx}{ds}, \quad \cos \beta = \frac{dy}{ds}, \quad \cos \gamma = \frac{dz}{ds}$$

69. **Circular Helix.** As an example of a space curve let us consider a line specified by the following kinematic conditions: a point M is in a uniform motion with constant velocity v_1 in a circle of radius a while the circle itself undergoes a uniform parallel displacement in the direction perpendicular to its plane with constant velocity v_2 . In this motion the point M describes a curve lying entirely on the corresponding circular cylinder (Fig. 93). This curve is known as a *screw line* or a *cylindrical* (or *circular*) *helix*. If the point M , when watched from the direction in which the circle is translated, is seen as rotating counterclockwise (clockwise), the curve is called a *right-hand* (*left-hand*) *screw line*.

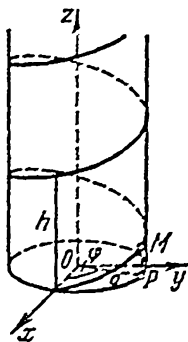


Fig. 93

Let us derive the equation of the screw line on condition that Oz is the axis of the cylinder, the parallel displacement is in the positive direction along Oz and the parameter t is the time of motion, the point M being in the position $(a, 0, 0)$ at the initial time moment $t=0$.

The angular velocity of the rotary motion of the point M being equal to $\frac{v_1}{a}$, the abscissa and the ordinate of M at the time moment t are, respectively, $x = a \cos \frac{v_1}{a} t$ and $y = a \sin \frac{v_1}{a} t$. The z -coordinate of the point M is equal to the altitude of the point above the xy -plane at time moment t , i.e., $z = v_2 t$. If, in place of time t , we take as the parameter the polar angle φ of the point P (which is the projection of the point M on the plane Oxy) then, since $\varphi = \frac{v_1}{a} t$, the equations of the screw line can be written in the form

$$x = a \cos \varphi, \quad y = a \sin \varphi, \quad z = c\varphi \quad \text{where } c = \frac{v_2}{v_1} a > 0$$

These equations specify a right-hand helix; the equations of a left-hand helix only differ from the above in the minus sign in the coefficient c .

After the angle φ is given an increment of 2π the point M returns to the original generator of the cylinder, the distance travelled along Oz being equal to $h = 2\pi c$. The latter quantity is called the *lead* or the *pitch of the helix*. Substituting $c = \frac{h}{2\pi}$ into the equations we arrive at the following form of parametric representation of

the helix, involving the pitch, convenient for practical purposes:

$$x = a \cos \varphi, \quad y = a \sin \varphi, \quad z = \frac{h}{2\pi} \varphi$$

Here both coefficients of the equations (a and h) have a simple geometrical meaning (see above).

A screw line possesses a number of interesting properties. We shall indicate two of them:

(1) The equation of the tangent to the screw line at a point $M_0(x_0, y_0, z_0)$ is written as

$$\frac{x-x_0}{-a \sin \varphi_0} = \frac{y-y_0}{a \cos \varphi_0} = \frac{z-z_0}{\frac{h}{2\pi}}$$

where φ_0 is the value of the angle φ corresponding to the point M_0 . It follows that

$$\cos \gamma = \frac{h}{2\pi \sqrt{a^2 + \left(\frac{h}{2\pi}\right)^2}} = \frac{h}{\sqrt{(2\pi a)^2 + h^2}}$$

which means that the cosine of the angle γ between the tangent and the z -axis, and therefore the angle γ itself, retain constant values at all the points of the screw line. The generators of the cylinder being parallel to Oz , the screw line intersects the generators of the cylinder at a constant angle dependent solely on the radius of the cylinder and on the pitch of the screw line.

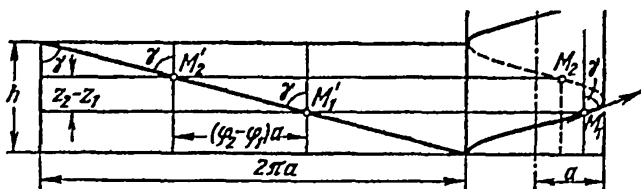


Fig. 94

(2) Let us fix a generator of the cylinder (on which the given screw line lies) in a plane passing through it and then develop (roll out) the cylinder upon that plane confining ourselves to a portion of the cylinder with height equal to the pitch of the screw line (such a portion contains one "coil" of the screw line; see Fig. 94). Under this transformation the base of the cylinder and its cross-sections parallel to the base go into parallel line segments of length $2\pi a$ while the generators go into the segments of length h perpendicular to the former segments. The development of the cylinder changing neither the angles between the lines lying on it nor the lengths of the lines, the screw line itself must go into a line intersecting the parallel line segments in the plane at a constant

angle. Such a line in the plane being only a straight line, we conclude that in the construction performed the screw line is transformed into the diagonal of the rectangle with sides $2\pi a$ and h . Fig. 94 demonstrates visually the above determined value of the cosine of the angle γ at which the screw line intersects the generators of the cylinder.

Let us take two arbitrary points M_1 and M_2 on the part of the screw line we have considered. The length of the arc of the screw line joining these points is equal to the length of the line segment M_1M_2' into which this arc goes as the cylinder is rolled out on the plane. Thus, *a screw line is a curve of shortest length between two points of a cylinder.*

A curve on an arbitrary surface through two its points having the shortest possible length compared with all the other curves connecting these points and lying on the surface is called a *geodesic* of that surface. For instance, the geodesics of a plane are the straight lines in that plane, and for a sphere the geodesics are the arcs of its great circles. Our investigation shows that a geodesic of a right circular cylinder is a screw line.

70. Vector Function of a Scalar Argument.

1. Vector function. Hodograph. It is known from vector algebra that any vector A whose projections on the coordinate axes are equal to x , y and z can be written in the form of the resolution along the coordinate axes

$$A = xi + yj + zk$$

where i , j and k are unit vectors of the coordinate axes. If the projections x , y , z are constant numbers we say that the vector A is *constant*. Now suppose that the projections of a vector are functions of a parameter t varying within a given interval:

$$x = x(t), \quad y = y(t), \quad z = z(t)$$

Then we say that the vector A itself is *variable*: to each value t of the parameter there corresponds a definite (vectorial) value of A :

$$A(t) = x(t)i + y(t)j + z(t)k$$

Definition. If to each value of a parameter t varying in a given interval there corresponds a definite vector $A(t)$ we call $A(t)$ a **vector function of the scalar argument t** .

As in vector algebra, we deal with free vectors; this means that two vectors having equal moduli and going in one direction or, which is the same, having equal projections on the coordinate axes, are identified and regarded as equal.

It is convenient to place the point of application of the given vector $A(t)$ at the origin; then, as t varies, the terminus of the

vector $\mathbf{A}(t)$ (having the coordinates $x(t)$, $y(t)$, $z(t)$) traces a curve L for which the relations

$$x = x(t), \quad y = y(t), \quad z = z(t)$$

serve as parametric equations.

The vector $\mathbf{A}(t)$ being nothing but the *radius vector* $\mathbf{r}$ of the moving point M of the curve L , we can specify this curve by means of the single *vector equation*

$$\mathbf{r} = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$$

Definition. The curve L described by the terminus of a variable vector $\mathbf{A}(t)$ issued from the origin is called the *hodograph* of the vector function $\mathbf{r} = \mathbf{A}(t)$.

The origin is then called the *pole of the hodograph*.

If the vector $\mathbf{A}(t)$ possesses a *variable modulus* while its direction remains unchanged its hodograph is a ray issued from the pole or a segment of such a ray. If the modulus of the vector is constant ($|\mathbf{A}(t)| = \text{const}$) while its *direction may vary* the hodograph is a curve lying on the sphere with centre at the pole of the hodograph and radius equal to the modulus of the vector $\mathbf{A}(t)$.

Vector functions of a scalar argument are most often encountered in kinematics when the motion of a material point is studied. The radius vector of a moving point is a function of time: $\mathbf{r} = \mathbf{A}(t)$; its hodograph is the *trajectory* of motion, and the equation $\mathbf{r} = \mathbf{A}(t)$ is the *equation of motion* (see Sec. 71).

II. The limit and the continuity of a vector function. The basic notions of mathematical analysis are defined for vector functions $\mathbf{A}(t)$ of a scalar argument by analogy with scalar functions.

Definition. A vector $\mathbf{B}$ is said to be the *limit of a vector function* $\mathbf{A}(t)$ as $t \rightarrow t_0$, which is written as

$$\lim_{t \rightarrow t_0} \mathbf{A}(t) = \mathbf{B}$$

if for all the values of t lying sufficiently close to t_0 the modulus of the difference of the vectors $|\mathbf{A}(t) - \mathbf{B}|$ becomes arbitrarily small.

If $\mathbf{A}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ and $\mathbf{B} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ then

$$|\mathbf{A}(t) - \mathbf{B}| = \sqrt{[x(t) - a]^2 + [y(t) - b]^2 + [z(t) - c]^2}$$

It is clear that the condition that $|\mathbf{A}(t) - \mathbf{B}|$ tends to zero as $t \rightarrow t_0$ implies $x(t) \rightarrow a$, $y(t) \rightarrow b$, $z(t) \rightarrow c$. The converse, obviously, is also true. Thus, we can briefly state that *the projections of the limit of a vector function $\mathbf{A}(t)$ are equal to the limits of its projections*.

Definition. A vector function $A(t)$ is said to be *continuous* for a given value t of the parameter if it is defined in a neighbourhood of the point t and if

$$\lim_{\Delta t \rightarrow 0} |A(t + \Delta t) - A(t)| = \lim_{\Delta t \rightarrow 0} |\Delta A(t)| = 0$$

The geometrical meaning of the difference $\Delta A(t)$ is clearly demonstrated in Fig. 95: in Fig. 95a it is equal to the vector

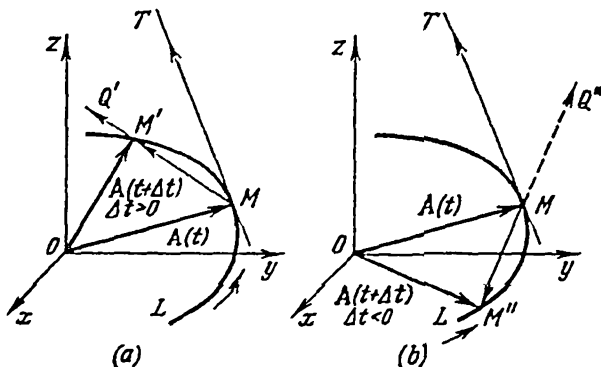


Fig. 95

$\overline{MM'}$ and in Fig. 95b to the vector $\overline{MM''}$. Let the resolution of the vector $A(t)$ into components along the coordinate axes be

$$A(t) = x(t)i + y(t)j + z(t)k \quad (*)$$

Then

$$A(t + \Delta t) = x(t + \Delta t)i + y(t + \Delta t)j + z(t + \Delta t)k$$

and, according to the rules of vector algebra,

$$\Delta A(t) = A(t + \Delta t) - A(t) = \Delta x i + \Delta y j + \Delta z k \quad (**)$$

where $\Delta x = x(t + \Delta t) - x(t)$, etc. Since

$$|\Delta A(t)| = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$$

the condition $|\Delta A(t)| \rightarrow 0$ implies that $\Delta x \rightarrow 0$, $\Delta y \rightarrow 0$, $\Delta z \rightarrow 0$. The converse is also obvious: if Δx , Δy , Δz tend to zero then $|\Delta A(t)|$ tends to zero as well. This means that the continuity of a vector function $A(t)$ is equivalent to the continuity of its projections $x(t)$, $y(t)$, $z(t)$. It is evident that the hodograph of a continuous vector function $r = A(t)$ is a *continuous* curve. In what follows we shall additionally suppose that this curve has a tangent at its every point (see Sec. 68).

III. **The derivative of a vector function.** To state the definition of the derivative of a vector function $\mathbf{A}(t)$ with respect to its scalar argument t we form the ratio

$$\frac{\Delta \mathbf{A}(t)}{\Delta t} = \frac{\mathbf{A}(t + \Delta t) - \mathbf{A}(t)}{\Delta t}$$

Geometrically, this ratio is represented by a vector going in the direction corresponding to the increase of t (this direction is indicated by the arrow in Fig. 95). For, if $\Delta t > 0$, the direction of the vector $\frac{\Delta \mathbf{A}(t)}{\Delta t} = \overline{MQ'}$ coincides with that of the vector $\Delta \mathbf{A}(t) = \overline{MM'}$, which means that it goes along the direction of the hodograph corresponding to the increase of the parameter t . If $\Delta t < 0$, the vector $\Delta \mathbf{A}(t) = \overline{MM''}$ is in the opposite direction, and, on dividing by the negative number Δt , we obtain the vector $\overline{MQ''}$ which is again in the direction of the increase of t .

Now we proceed to consider the limit of the ratio $\frac{\Delta \mathbf{A}(t)}{\Delta t}$ as $\Delta t \rightarrow 0$.

Definition. *The derivative of a vector function $\mathbf{A}(t)$ is the limit (provided it exists)*

$$\lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{A}(t)}{\Delta t} = \mathbf{A}'(t) = \frac{d\mathbf{A}(t)}{dt}$$

By the definition of the limit (see II), the derivative of a vector function is itself a *vector*. It is denoted by $\mathbf{A}'(t)$ or $\frac{d\mathbf{A}(t)}{dt}$. As for the direction of the vector $\mathbf{A}'(t)$, observe that (see Fig. 95), as $\Delta t \rightarrow 0$, the point M' (or M'') tends to the point M , and therefore the secant MM' (or MM'') approaches the *tangent* at the point M . Thus, *the derivative $\mathbf{A}'(t)$ is the vector $\overline{MT}$ tangent to the hodograph of the vector function $\mathbf{A}(t)$ and going in the direction corresponding to the increase of the parameter t .*

If the modulus of a vector function $\mathbf{A}(t)$ is constant (whereas its direction may vary) its derivative $\mathbf{A}'(t)$ is a *vector perpendicular to the vector $\mathbf{A}(t)$* . Indeed, in this case the hodograph lies on a sphere, and therefore the derivative $\mathbf{A}'(t)$ as vector tangent to the hodograph is perpendicular to the radius vector $\mathbf{A}(t)$.

Thus, the derivative of a vector of constant modulus is perpendicular to that vector.

Now we proceed to the practical determination of the derivative $\mathbf{A}'(t)$ of a given vector function $\mathbf{A}(t)$. Let the vector function $\mathbf{A}(t)$ be specified by its resolution

$$\mathbf{A}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$$

Then, according to (**), we have

$$\frac{\Delta \mathbf{A}(t)}{\Delta t} = \frac{\Delta x}{\Delta t} \mathbf{i} + \frac{\Delta y}{\Delta t} \mathbf{j} + \frac{\Delta z}{\Delta t} \mathbf{k}$$

On passing to the limit as $\Delta t \rightarrow 0$ and using the rule for finding the limit of a vector function we arrive at the resolution of the derivative $\mathbf{A}'(t)$ along the unit vectors of the coordinate axes:

$$\mathbf{A}'(t) = x'(t) \mathbf{i} + y'(t) \mathbf{j} + z'(t) \mathbf{k}$$

It follows that

$$|\mathbf{A}'(t)| = \sqrt{x'^2(t) + y'^2(t) + z'^2(t)}$$

Now, recalling the expression for the differential of arc length ds (Sec. 68), we can rewrite the latter equality in the form

$$|\mathbf{A}'(t)| = \frac{ds}{dt}$$

Hence, the modulus of the derivative of a vector function $|\mathbf{A}'(t)|$ is equal to the derivative of the arc length of the hodograph with respect to the argument t . It should be stressed that the modulus of the derivative $|\mathbf{A}'(t)|$ is not equal to the derivative of the modulus $(|\mathbf{A}(t)|)'$. This can be illustrated by the example of a variable vector of constant length: the modulus of such a vector is a constant number and its derivative is simply equal to zero whereas the derivative of the vector itself is a new vector perpendicular to the former.

Taking advantage of the expression of the derivative $\mathbf{A}'(t)$ we can easily show that all the basic differentiation rules for scalar functions are extended almost without any changes to vector functions:

- (1) $[\mathbf{A}_1(t) + \mathbf{A}_2(t)]' = \mathbf{A}'_1(t) + \mathbf{A}'_2(t);$
- (2) $[f(t) \mathbf{A}(t)]' = f'(t) \mathbf{A}(t) + f(t) \mathbf{A}'(t)$

where $f(t)$ is a scalar function; in particular, $[C\mathbf{A}(t)]' = C\mathbf{A}'(t)$ where C is a constant scalar.

The differentiation rules for the scalar and vector products $\mathbf{A}_1(t) \cdot \mathbf{A}_2(t)$ and $\mathbf{A}'_1(t) \times \mathbf{A}'_2(t)$ of two vector functions $\mathbf{A}_1(t)$ and $\mathbf{A}_2(t)$ are also completely analogous to the corresponding rules for the product of scalar functions:

- (3) $[\mathbf{A}_1(t) \cdot \mathbf{A}_2(t)]' = \mathbf{A}'_1(t) \cdot \mathbf{A}_2(t) + \mathbf{A}_1(t) \cdot \mathbf{A}'_2(t);$
- (4) $[\mathbf{A}_1(t) \times \mathbf{A}_2(t)]' = [\mathbf{A}'_1(t) \times \mathbf{A}_2(t)] + [\mathbf{A}_1(t) \times \mathbf{A}'_2(t)].$

(In Rule 4 it is important to retain the order of the vectorial factors.)

The proof of all these rules is left to the reader.

By means of successive differentiations we readily find the derivatives of higher orders of a vector function. For example,

$$\mathbf{A}''(t) = x''(t)\mathbf{i} + y''(t)\mathbf{j} + z''(t)\mathbf{k}$$

and so on.

71*. Applications to Mechanics. Differentiation of vector functions of a scalar argument (see the foregoing section) has an important application to particle mechanics in the investigation of the velocity and of the acceleration of a *curvilinear* motion of a point. (In Secs. 38, I and 54, I we considered the problem of determining the velocity and the acceleration of a *rectilinear* motion of a point.)

Let t be time of motion, and let the hodograph of a vector function $\mathbf{r} = \mathbf{A}(t)$ be the trajectory of a moving point M . We shall denote by s the distance from a fixed initial point of the trajectory to the moving point M reckoned along the trajectory and taken with the sign $+$ or $-$ depending on whether the point M is moved from the initial position in the positive or negative direction chosen (in an arbitrary way) along the trajectory. The location of the point M on the trajectory is then completely specified by the magnitude s which is a *curvilinear coordinate* of the point M . The equation $s = s(t)$ expresses the *law of motion along the trajectory*.

By definition, *the velocity of the point M at a given time moment t is the derivative of the vector function $\mathbf{r} = \mathbf{A}(t)$ with respect to time:*

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \mathbf{A}'(t)$$

Consequently, *the velocity of the moving point is a tangent vector to the trajectory at the corresponding point going in the direction of motion* (that is, in the direction corresponding to the increase of t).

The *modulus of the velocity* is expressed (see Sec. 70, III) by the formula

$$|\mathbf{v}| = |\mathbf{A}'(t)| = \frac{ds}{dt}$$

that is, it is equal to the *derivative of the curvilinear coordinate s with respect to time t* .

If the motion is rectilinear the scalar magnitude $\frac{ds}{dt}$ completely specifies the velocity. It is this quantity that we called the velocity of rectilinear motion of a point.

The vector

$$\mathbf{w} = \frac{d\mathbf{v}}{dt} = \frac{d^2\mathbf{r}}{dt^2}$$

is called the *acceleration* of motion. We shall transform this expression of the acceleration w . To this end, let us denote by τ_1 unit tangent vector to the hodograph (Fig. 96). Then

$$v = |v| \cdot \tau_1 = \frac{ds}{dt} \tau_1$$

On differentiating this product (see Rule 2) we obtain

$$w = \frac{d^2s}{dt^2} \tau_1 + \frac{ds}{dt} \frac{d\tau_1}{dt}$$

Now let us consider the derivative $\frac{d\tau_1}{dt}$. Since τ_1 is a unit vector ($|\tau_1| = 1$) its derivative is perpendicular to it, i.e. it goes along a *normal* to the curve (this particular normal is called the *principal normal* of the curve). Denoting unit vector of the principal normal by ν_1 we can write

$$\frac{d\tau_1}{dt} = \left| \frac{d\tau_1}{dt} \right| \nu_1$$

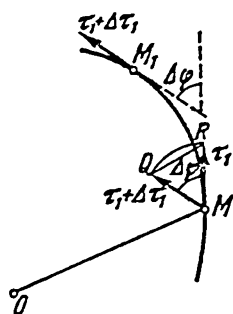


Fig. 96

Fig. 96 indicates that $|\Delta\tau_1|$ is the length of the chord subtending the arc $\widehat{MM_1}$ of the circle with radius $|\Delta\tau_1|$, $\Delta\varphi$ being a central angle of this circle equal to the angle of contingence of the arc $\widehat{MM_1}$ of the hodograph (see Sec. 67). By the equivalence of an infinitesimal arc and its chord, we have

$$\left| \frac{d\tau_1}{dt} \right| = \left| \frac{d\varphi}{dt} \right|$$

Next we transform the expression of $\frac{d\varphi}{dt}$:

$$\left| \frac{d\varphi}{dt} \right| = \left| \frac{d\varphi}{ds} \frac{ds}{dt} \right| = \frac{ds}{dt} K$$

where $K = \left| \frac{d\varphi}{ds} \right|$ is the limit of the ratio of the angle of contingence to the arc length. As in the case of a plane curve, the magnitude K is referred to as the *curvature* of the space curve $r = A(t)$, and its reciprocal $\rho = \frac{1}{K}$ is the *radius of curvature*¹. The substitution of the expression of $\frac{d\tau_1}{dt}$ into the formula expressing the acceleration w yields

$$w = \frac{d^2s}{dt^2} \tau_1 + \frac{\left(\frac{ds}{dt} \right)^2}{\rho} \nu_1$$

¹ In our course we do not derive the formula for the curvature K of a space curve.

The first component on the right-hand side $\mathbf{w}_t = \frac{d^2s}{dt^2} \boldsymbol{\tau}_1$ is called the *tangential acceleration*; its modulus is equal to the second derivative of the curvilinear coordinate with respect to time, and its direction coincides with that of the tangent to the trajectory of

motion. The second component $\mathbf{w}_n = \frac{\left(\frac{ds}{dt}\right)^2}{\rho} \mathbf{v}_1 = \frac{|v|^2}{\rho} \mathbf{v}_1$ is called the *normal* (or *centripetal*) *acceleration*; it goes along the principal normal to the trajectory, its modulus being equal to the square of the magnitude of the velocity divided by the radius of curvature.

The plane passing through the tangent and the principal normal at a given point of a curve is called the *osculating plane*. In the general case, the osculating plane of a space curve varies from point to point. For a plane curve the osculating plane is common for all its points and coincides with the plane of the curve.

The latter formula for the acceleration shows that the vector $\mathbf{w}$ always lies in the osculating plane.

If the motion is *rectilinear* we have $\mathbf{w}_n = 0$, and the acceleration is completely specified by the scalar magnitude $\frac{d^2s}{dt^2}$.

If the motion is *uniform*, that is if the absolute value of the velocity $|v| = \frac{ds}{dt}$ is constant, the component $\mathbf{w}_t$ of the acceleration turns into zero, and only the centripetal acceleration

$$\mathbf{w}_n = \frac{|v|^2}{\rho} \mathbf{v}_1$$

can be different from zero. In particular, if a point moves uniformly in a circle of radius R , we have $\rho = R$ and $|\mathbf{w}_n| = \frac{|v|^2}{R}$.

In this case the acceleration $\mathbf{w}$ is directed toward the centre of the circle (in the general case of an arbitrary curvilinear motion the component $\mathbf{w}_n$ is directed toward the centre of curvature which is defined for space curves by analogy with plane curves; the above consideration and the latter fact account for the term "centripetal acceleration").

§ 6. Complex Functions of a Real Argument

72. Complex Numbers. We begin with a brief review of the basic properties of complex numbers and the rules for arithmetical operations on them which are familiar to the reader from elementary mathematics.

The number

$$z = x + iy$$

where x and y are any real numbers and i the so-called *imaginary unit* ($i^2 = -1$) is called a *complex number*, x being its *real part* and y the *imaginary part*. By definition, two complex numbers are equal if and only if their real and imaginary parts coincide. Thus, the equality

$$x_1 + iy_1 = x_2 + iy_2$$

is equivalent to the two equalities

$$x_1 = x_2 \quad \text{and} \quad y_1 = y_2$$

The equality $z = x + iy = 0$ is equivalent to the conditions $x = 0$ and $y = 0$.

The complex number $z = x + iy$ can be represented by a point in the xy -plane (Fig. 97) whose abscissa is the real part of z and ordinate the imaginary part of z . In this representation the axis

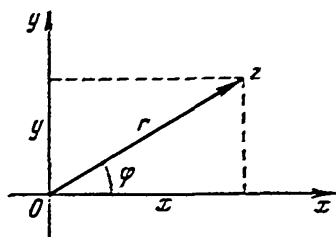


Fig. 97

of abscissas is called the *real axis* (or the *axis of reals*) and the axis of ordinates the *imaginary axis* (the *axis of imaginaries*), the plane Oxy itself being referred to as the *complex plane*. The representing point of a complex number z is referred to as the point z .

A complex number can also be regarded as represented by the vector applied to the origin with terminus at the representing point of that complex number; the projections of this vector on the coordinate axes x and y are, respectively, the real and the imaginary parts of the complex number z . This geometrical interpretation of complex numbers as vectors will be used in the following section.

If $y = 0$ the complex number $z = x + i0 = x$ is a real number represented by the corresponding point of the real axis; if $x = 0$ the number $z = 0 + iy = y$ is called *pure imaginary* and is represented by the point $(0, y)$ lying on the imaginary axis.

The location of the point representing a complex number z can also be specified by the polar coordinates r and φ (Fig. 97). In this specification the magnitude $r = \sqrt{x^2 + y^2}$ is called the *modulus* (or the *absolute value*) and φ the *argument* of the complex number z ; they are respectively denoted as

$$r = |z| \quad \text{and} \quad \varphi = \text{Arg } z$$

The magnitude $\text{Arg } z$ is defined to within an integral multiple of 2π , i.e. $2k\pi$ where k is an integer. If the angle φ is chosen

between the limits $-\pi$ and π it is called the *principal value* of the argument and is denoted as $\arg z$.

If $z = x + iy$ then $x - iy$ is termed the *conjugate complex number* of z and is denoted by $\bar{z}$. The points $\bar{z}$ and z are obviously symmetric with respect to the real axis (i.e. to the axis of abscissas). It is apparent that $|\bar{z}| = |z|$ and $\arg \bar{z} = -\arg z$.

Using the transformation formulas from Cartesian coordinates to polar coordinates ($x = r \cos \varphi$, $y = r \sin \varphi$) we can represent any complex number different from zero in the so-called *trigonometric form*:

$$z = x + iy = r(\cos \varphi + i \sin \varphi)$$

The arithmetical operations on complex numbers are carried out according to the following rules:

$$(1) (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$

that is, the sum of two complex numbers is a complex number whose real and imaginary parts are, respectively, the sum of the real parts and the sum of the imaginary parts of the summands. In the geometrical representation this rule corresponds to the well-known rule of vector algebra for the addition of vectors.

$$(2) (x_1 + iy_1) - (x_2 + iy_2) = (x_1 - x_2) + i(y_1 - y_2)$$

This means that, as usual, the subtraction is regarded as an operation inverse to the addition, and the relation written here simply expresses analytically the ordinary rule for the subtraction of vectors. The difference of two complex numbers $z_1 - z_2$ is represented by the vector with origin at the point z_2 and terminus at the point z_1 . Hence, $|z_1 - z_2|$ is the length of the line segment joining the points z_1 and z_2 .

$$(3) (x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)$$

Taking the factors in the trigonometric form we readily derive the relation

$$\begin{aligned} r_1(\cos \varphi_1 + i \sin \varphi_1) r_2(\cos \varphi_2 + i \sin \varphi_2) &= \\ &= r_1 r_2 [\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)] \end{aligned}$$

Thus, as complex numbers are multiplied, their moduli are also multiplied while the arguments are added together.

This rule is readily extended, in succession, to an arbitrary number of factors. If there are n equal factors we obtain the rule for raising a complex number to the n th power:

$$[r(\cos \varphi + i \sin \varphi)]^n = r^n (\cos n\varphi + i \sin n\varphi)$$

where n is a positive integer. This result is known as *De Moivre's formula*¹.

¹ De Moivre, A. (1667-1754), an English mathematician.

The product of a complex number z by its conjugate $\bar{z}$ is a positive real number equal to the square of the modulus of z (the moduli of z and $\bar{z}$ are equal): $z\bar{z} = (x+iy)(x-iy) = x^2 + y^2 = |z|^2 = |\bar{z}|^2$.

$$(4) \frac{x_1 + iy_1}{x_2 + iy_2} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{(x_2 + iy_2)(x_2 - iy_2)} = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i \frac{-x_1y_2 + x_2y_1}{x_2^2 + y_2^2}$$

on condition that $x_2 + iy_2 \neq 0$. The division is an operation inverse to the multiplication. Taking the dividend and the divisor in the trigonometric form we obtain

$$\frac{r_1(\cos \varphi_1 + i \sin \varphi_1)}{r_2(\cos \varphi_2 + i \sin \varphi_2)} = \frac{r_1}{r_2} [\cos(\varphi_1 - \varphi_2) + i \sin(\varphi_1 - \varphi_2)]$$

Consequently, as a complex number is divided by another (nonzero) complex number, the modulus of the dividend is divided by that of the divisor and the argument of the divisor is subtracted from the argument of the dividend.

The analysis of the formulas presented here shows that the rules expressed by them can be stated in the general form as

the arithmetical operations on complex numbers are performed according to the ordinary rules for operations on the binomials $(x+iy)$ after which i^2 is everywhere replaced by -1 .

The general laws governing the arithmetical operations on real numbers and the properties of these operations are extended to complex numbers without any changes. It should be noted that if all the complex numbers involved in an arithmetical operation are replaced by their conjugates the result of the operation is the complex conjugate of the former result. Indeed, if the signs of y_1 and y_2 are changed to the opposite the signs of the right-hand sides of all the four formulas (1), (2), (3) and (4) also change. This property taking place in each of the four basic arithmetical operations, the same applies to its arbitrary combinations.

73. Definition of a Complex Function. Differentiation.

I. Suppose that we are given a vector function of a scalar argument whose projection on the z -axis is identically equal to zero for all the values of the parameter t . Then

$$\mathbf{A}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} \quad (*)$$

and the curve $\mathbf{r} = \mathbf{A}(t)$ (the hodograph of this function) entirely lies in the plane Oxy . In this case it is convenient to consider the vector $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ as representing geometrically the complex number $z = x + iy$ and, accordingly, to speak, instead of the vector function $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, of the complex function $z(t) = x(t) + iy(t)$ of the real variable t .¹

¹ Complex functions of a complex argument are studied in the theory of functions of a complex variable (e.g. see [3], [16], [19]).

Definition. If with each value of a real parameter t there is associated a definite complex number

$$z(t) = x(t) + iy(t) \quad (**)$$

where $x(t)$ and $y(t)$ are functions assuming real values, $z(t)$ is called a **complex function** of the real argument t .

As a rule, the parameter t ranges within an interval (finite or infinite).

The **hodograph** of a complex function $z(t) = x(t) + iy(t)$ is, by definition, the curve with parametric equations $x = x(t)$, $y = y(t)$; thus, the hodographs of the vector function (*) and of the complex function (**) coincide.

The definitions of the limit and of the continuity of a complex function of a real argument are completely analogous to the corresponding definitions for a vector function, and we leave it to the reader to state them. We only note that the continuity of a complex function $z(t) = x(t) + iy(t)$ is equivalent to the continuity of its real and imaginary parts $x(t)$ and $y(t)$. The hodograph of a continuous function $z(t)$ traced as the parameter t varies from t_1 to t_2 is a continuous line joining the points $z(t_1)$ and $z(t_2)$ in the complex plane.

Examples. (1) For the function $z(t) = t + it^2$ ($-\infty < t < +\infty$) we have $x = t$ and $y = t^2$. The hodograph is the parabola $y = x^2$. When t varies from $-\infty$ to $+\infty$ the moving point of the parabola traverses it so that the upper (infinite) region bounded by the parabola always remains on the left.

(2) Let us take the function $z(t) = \cos t + i \sin t$ ($0 \leq t \leq 2\pi$). Since $|z(t)| = \sqrt{\cos^2 t + \sin^2 t} = 1$, the hodograph of this complex function of the parameter t is the circle of radius 1 with centre at the origin. As t increases from 0 to 2π the variable point describes the circle once in the counterclockwise direction. If the parameter t only varied within the interval $[0, \pi]$ the hodograph would be the upper half-circle with initial point $z = 1$ and terminal point $z = -1$.

(3) The hodograph of the complex function $z(t) = 1 + i \cos t$ ($0 \leq t \leq 2\pi$) is the line segment parallel to the axis of imaginaries joining the point $z_1 = 1 + i$ to the point $z_2 = 1 - i$; as t runs through the given interval this line segment is described twice; first downward and then upward.

II. The derivative of a complex function of a real variable.

The derivative of a complex function $z(t)$ is defined in the ordinary way, that is, as the limit of the ratio of the increment of the function $\Delta z = z(t + \Delta t) - z(t)$ to the increment of the independent variable Δt :

$$z'(t) = \lim_{\Delta t \rightarrow 0} \frac{\Delta z}{\Delta t}$$

In its turn, the derivative $z'(t)$ is a complex function of the same argument. In the geometrical interpretation, the representing vector of the complex number $z'(t_0)$ is parallel to the tangent line to the hodograph of the function $z(t)$ at the point of the hodograph corresponding to the value $t=t_0$ of the parameter.

Repeating without changes the argument used in the derivation of the formula for the derivative of an arbitrary vector function of a scalar argument we obtain for a given complex function $z(t)=x(t)+iy(t)$ the differentiation formula

$$z'(t)=x'(t)+iy'(t)$$

This formula indicates that a complex function $z(t)=x(t)+iy(t)$ can be differentiated as an ordinary linear combination in which i is regarded as an ordinary constant factor. This simple rule is of course justified only after the definitions discussed in this section have been introduced. It can also be easily verified that all the differentiation rules for real functions are generalized without any changes to complex functions.

In conclusion we note that the product of complex functions is computed according to the multiplication rule for complex numbers.

74. Exponential Function. Euler's Formulas. In this section we state the definition of the exponential function with imaginary exponent $z=e^{it}$ and study its properties. This function plays an extremely important role both in mathematics and in its applications to various applied sciences such as electrical engineering, the theory of vibrations and the like.

Definition. The *complex exponential function* e^{it} with imaginary exponent is specified, by definition, by the relation

$$e^{it} = \cos t + i \sin t \quad (*)$$

where the parameter t may assume any real values.

Relation (*) is known as *Euler's formula*.

The justification of this definition (which may at first glance seem obscure¹) lies in the fact that all the operation rules for the ordinary exponential function $y=e^x$ (including the differentiation formula) remain valid for the function e^{it} thus defined.

Before checking these rules let us dwell on some properties of the function e^{it} directly implied by its definition. From the result established in Example 2 of Sec. 73 it follows that the modulus of the function e^{it} is equal to unity; this is also implied by the fact that the right-hand side of formula (*) is nothing but the trigonometric form of the complex number with modulus equal to 1 and argument equal to t :

$$|e^{it}| = 1, \text{ Arg } e^{it} = t$$

¹ Later on (Chapter XI) we shall arrive at the same definition on the basis of some completely different considerations.

The hodograph of the function $z=e^{it}$ is unit circle with centre at the origin (see Fig. 98). For $t=0$ we obtain the point $z=1$, for $t=\frac{\pi}{2}$ the point $z=i$, for $t=\pi$, the point $z=-1$, etc. Thus,

$$e^{i\frac{\pi}{2}}=i, \quad e^{i\pi}=-1, \quad e^{i\frac{3\pi}{2}}=-i$$

As the parameter t increases from 0 to 2π the variable point $z=e^{it}$ moves in the unit circle counterclockwise and returns to the original position $z=1$ for $t=2\pi$:

$$e^{2\pi i}=\cos 2\pi+i\sin 2\pi=1$$

If t is given further increase to run from 2π to 4π the point e^{it} again traverses the circle and so on. Therefore it becomes clear that *the function $z=e^{it}$ is periodic with period $T=2\pi$* . For, we have

$$\begin{aligned} e^{i(t+2\pi)} &= \cos(t+2\pi) + i\sin(t+2\pi) = \\ &= \cos t + i\sin t = e^{it} \end{aligned}$$

Note that if n is an integer then

$$e^{2n\pi i}=1 \text{ and } e^{(2n+1)\pi i}=-1$$

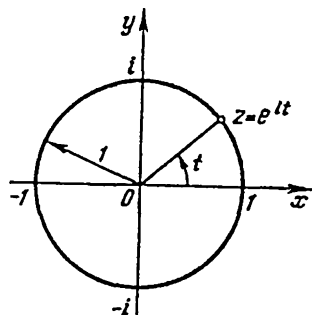


Fig. 98

The basic operation rule for the real exponential function states that if x_1 and x_2 are two real numbers then $e^{x_1}e^{x_2}=e^{x_1+x_2}$. Let us check that this rule also remains valid in the case of imaginary exponents. Indeed, according to the trigonometric form of the multiplication rule for complex functions, we can write

$$\begin{aligned} e^{it_1}e^{it_2} &= (\cos t_1 + i\sin t_1)(\cos t_2 + i\sin t_2) = \cos(t_1+t_2) + \\ &+ i\sin(t_1+t_2) = e^{i(t_1+t_2)} \end{aligned}$$

where t_1 and t_2 are any real numbers (positive or negative or zero).

Euler's formula allows us to write complex numbers in the so-called *exponential form*: if $|z|=r$ and $\text{Arg } z=\varphi$ then

$$z=r(\cos \varphi + i\sin \varphi)=re^{i\varphi}$$

The multiplication rule shows that if $z_1=r_1e^{i\varphi_1}$ and $z_2=r_2e^{i\varphi_2}$ then

$$z_1z_2=r_1e^{i\varphi_1}r_2e^{i\varphi_2}=r_1r_2e^{i(\varphi_1+\varphi_2)}$$

The product of a complex number $z=re^{i\varphi}$ by the factor $e^{i\alpha}$ is $ze^{i\alpha}=re^{i(\varphi+\alpha)}$. The geometrical interpretation of this fact is that the multiplication by $e^{i\alpha}$ makes the vector representing the complex number z rotate about the origin through the angle α . In particular, putting $\alpha=\frac{\pi}{2}$ we see that the multiplication by $e^{i\frac{\pi}{2}}=i$

results in the rotation of the representing vector of the number z through 90° in the counterclockwise direction.

The natural requirement that the multiplication rule should be extended to complex functions leads to the following definition of the exponential function with an arbitrary *complex* exponent $\alpha + i\beta$ whose real and imaginary parts α and β are arbitrary real numbers:

$$e^{(\alpha + i\beta)t} = e^{\alpha t} e^{i\beta t} = e^{\alpha t} (\cos \beta t + i \sin \beta t) \quad (**)$$

This relation implies

$$|e^{(\alpha + i\beta)t}| = e^{\alpha t}, \quad \text{Arg } e^{(\alpha + i\beta)t} = \beta t$$

For instance, the number e to the power $2 + 3i$ is found thus: $e^{2+3i} = e^2 (\cos 3 + i \sin 3)$.

Now we proceed to differentiate the function e^{it} . The differentiation rule for complex functions of a real independent variable implies

$$(e^{it})' = -\sin t + i \cos t = i (\cos t + i \sin t) = i e^{it}$$

Thus, the complex function e^{it} is differentiated as if i were an ordinary constant coefficient.

The geometrical representation of the differentiation of e^{it} reduces to the rotation of the vector $z = e^{it}$ through 90° in the counterclockwise direction.

Generally, the application of formula (**) shows that

$$[e^{(\alpha + i\beta)t}]' = (\alpha + i\beta) e^{(\alpha + i\beta)t}$$

This means that the formula

$$(e^{\lambda t})' = \lambda e^{\lambda t}$$

is valid in all cases, that is for any (real or complex) number λ . Euler's formula

$$e^{it} = \cos t + i \sin t$$

expresses the exponential function with imaginary exponent in terms of the trigonometric functions sine and cosine. Substituting $-t$ for t into this formula we obtain

$$e^{-it} = \cos(-t) + i \sin(-t) = \cos t - i \sin t$$

This shows that e^{it} and e^{-it} are two conjugate complex numbers.

Now we readily express the trigonometric functions $\cos t$ and $\sin t$ in terms of the exponential functions e^{it} and e^{-it} :

$$\cos t = \frac{e^{it} + e^{-it}}{2}, \quad \sin t = \frac{e^{it} - e^{-it}}{2i} \quad (***)$$

The last two relations are also referred to as Euler's formulas.

§ 7. Solution of Equations

75. General Properties of Equations. The need for solving equations is encountered in various problems both of purely mathematical and of applied character. For instance, when investigating a function $y=f(x)$ we should solve the equations $f(x)=0$, $f'(x)=0$ and $f''(x)=0$ in order to determine, respectively, the zeros, the points of extremum and the abscissas of the points of inflection of the function. These equations by far not always admit of a precise solution, and therefore their roots are often found *approximately*. We see that such questions (and many other problems) require approximate methods of solving equations. Before proceeding to study them we shall discuss general properties of equations some of which are already familiar to the reader.

I. Algebraic equations. Factorization of polynomials. Given a polynomial

$$P(x) = a_0x^n + a_1x^{n-1} + \dots + a_n$$

with real or complex coefficients ($a_0 \neq 0$), the equation

$$a_0x^n + a_1x^{n-1} + \dots + a_n = 0 \quad (*)$$

is referred to as an *algebraic equation of the n th degree*.

The fundamental theorem of algebra asserts that *any algebraic equation of degree $n > 0$ has at least one (real or complex) root*. But this theorem does not provide any practical means for finding the root and only guarantees its existence (the proof of this theorem is far beyond the framework of our course and we do not present it here).

On dividing a given polynomial of the n th degree $P(x)$ by an arbitrary binomial $x-\alpha$ we obtain a polynomial of the $(n-1)$ th degree $Q(x)$ in the quotient and a number R in the remainder. This can be written as the identity

$$P(x) = (x-\alpha) Q(x) + R$$

The way this identity has been derived does not allow us to resort to it for $x=\alpha$ since we cannot divide by zero. But we can pass to the limit in it as $x \rightarrow \alpha$, which results in $P(\alpha)=R$. Thus, *the remainder appearing in the division of a polynomial $P(x)$ by a binomial $x-\alpha$ is equal to the value of that polynomial at the point $x=\alpha$ (this proposition is known as Bézout's theorem¹).*

If α is a root of the equation $P(x)=0$ then $R=P(\alpha)=0$ and

$$P(x) = (x-\alpha) Q(x)$$

Let x_1 be a root of equation (*); then

$$P(x) = (x-x_1) Q_1(x)$$

¹ Bézout, Ê (1730-1783), a French mathematician.

where $Q_1(x)$ is a polynomial of the $(n-1)$ th degree. It is readily observed that its *leading coefficient* (i.e., the coefficient in x^{n-1}) is equal to that of the polynomial $P(x)$, that is, to a_0 .

If the degree of $Q_1(x)$ is different from zero, that is, if this polynomial is not identically equal to the constant a_0 , the fundamental theorem can be applied repeatedly. Let x_2 be a root of $Q_1(x)$; then $Q_1(x) = (x-x_2) Q_2(x)$ where the degree of the polynomial $Q_2(x)$ is equal to $n-2$, its leading coefficient again being a_0 . The expression for $P(x)$ now takes the form

$$P(x) = (x-x_1)(x-x_2) Q_2(x)$$

On repeating this process n times we arrive at the factorization of the polynomial $P(x)$ into linear factors:

$$P(x) = a_0(x-x_1)(x-x_2) \dots (x-x_n)$$

In this sequence of similar operations there may appear coincident roots among $x_1, x_2, \dots, x_n$. If a root in this sequence is repeated k times it is spoken of as a *k-fold root* or a root of *multiplicity k*. If $k=1$ the root occurs only once and is called *simple*.

Let the root x_1 have multiplicity k_1 , the root x_2 multiplicity k_2 , etc. Then the factorization can be rewritten in the form

$$P(x) = a_0(x-x_1)^{k_1}(x-x_2)^{k_2} \dots (x-x_r)^{k_r} \quad (**)$$

where r is the number of pairwise distinct roots, and the sum of the multiplicities of all the roots is equal to n :

$$k_1 + k_2 + \dots + k_n = n$$

The latter relation can be read as *every algebraic equation of the n th degree has exactly n roots on condition that each of the pairwise distinct roots is counted as many times as its multiplicity*.

What has been said applies to equations (*) with arbitrary complex coefficients. Let us now suppose that all the coefficients $a_0, a_1, \dots, a_n$ are *real* numbers. Then the complex roots of equation (*) (provided that they exist) are grouped into pairs of conjugate complex numbers. This property is a consequence of the remark at the end of Sec. 72. For, if the coefficients of the given polynomial $P(x)$ are real numbers its values computed for two conjugate complex values of x must be a pair of conjugate complex numbers ($P(\alpha-i\beta) = \overline{P(\alpha+i\beta)}$), and it follows that if $P(\alpha+i\beta)=0$ then also $P(\alpha-i\beta)=0$. It can readily be shown that any two conjugate complex roots $\alpha+i\beta$ and $\alpha-i\beta$ have the same multiplicity. By the way, this implies that *any polynomial of odd degree with real coefficients possesses at least one real root*.

Combining in factorization (**) the factors corresponding to a chosen pair of conjugate complex roots $\alpha \pm i\beta$ of multiplicity l

and performing the multiplication we receive

$$(x - \alpha - i\beta)^l (x - \alpha + i\beta)^l = (x^2 - 2\alpha x + \alpha^2 + \beta^2)^l = (x^2 + px + q)^l$$

where $p = -2\alpha$ and $q = \alpha^2 + \beta^2$. The discriminant $p^2 - 4q$ of this quadratic trinomial is of course negative since its roots are complex.

On performing the same transformation for all the factors containing conjugate complex roots we arrive at the factorization of the polynomial $P(x)$ into *real* linear and quadratic factors:

$$P(x) = a_0 (x - x_1)^{k_1} \dots (x^2 + p_1 x + q_1)^{l_1} \dots$$

where k_i 's are the multiplicities of the real roots and l_i 's of the conjugate complex roots. It is apparent that $(k_1 + \dots) + 2(l_1 + \dots) = n$.

These factorizations will be used in the next chapter.

This analysis implies some simple consequences important for our further aims.

1. If $P(x)$ is a polynomial identically equal to zero for any values of x , that is, $P(x) \equiv 0$, all its coefficients are zero:

$$a_0 = a_1 = \dots = a_n = 0$$

For, if we had $a_0 \neq 0$ the polynomial $P(x)$ would possess not more than n distinct roots, if we had $a_0 = 0$ and $a_1 \neq 0$ the number of distinct roots would not exceed $n-1$ and so on whereas the condition $P(x) \equiv 0$ means that the number of such roots is infinite. It also follows that if a polynomial of the n th degree has more than n roots it is identically equal to zero.

2. If the values of two polynomials coincide for any values of x , that is

$$a_0 x^n + a_1 x^{n-1} + \dots + a_n \equiv b_0 x^m + b_1 x^{m-1} + \dots + b_m$$

then their degrees and their coefficients in like powers of x coincide. Indeed, on transposing all the terms of the identity to the left-hand side we obtain (for definiteness, it is assumed that $n > m$):

$$\begin{aligned} a_0 x^n + \dots + a_{n-m-1} x^{m+1} + (a_{n-m} - b_0) x^m + \dots \\ \dots + (a_{n-1} - b_{m-1}) x + (a_n - b_m) \equiv 0 \end{aligned}$$

Now it follows from the first consequence that, on the one hand, $a_0 = a_1 = \dots = a_{n-m-1} = 0$, i.e. the leading term of the first polynomial is $a_{n-m} x^m$, and, on the other hand, $a_{n-m} = b_0$, ..., $a_{n-1} = b_{m-1}$, $a_n = b_m$. These two conditions show that the polynomials coincide identically.

There are two simple cases in which an equation with arbitrary coefficients is readily solved: $n=1$ (a linear equation) and $n=2$ (a quadratic equation), both cases being studied in elementary mathematics. For the cases $n=3$ and $n=4$ (accordingly, equations of degrees 3 and 4 called *cubics* and *quartics*) there also exist general formulas for the roots (known as *Cardano's solution of*

the cubic and Ferrari's solution of the quartic¹) but they are extremely lengthy and complicated, and it is difficult to use them in practice. As for the equations of the fifth degree and higher, it is proved that there exist no general formulas expressing their roots in terms of the coefficients of the equations by means of a finite number of arithmetical operations and operations of extracting roots (in this connection it is said that in the general case such equations are not *solvable by radicals*²). Some concrete examples of equations that are not solvable by radicals are also known. There are only some particular cases (most of which are treated in elementary mathematics) when the solution by radicals is possible (as a rule, this is the case when the left-hand side of the equation can readily be written as a product of polynomials of lower degrees).

This discussion once again indicates the importance of approximate methods of solution.

II. An equation $R(x)=0$ where $R(x)$ is a *rational fraction* of the form $\frac{P(x)}{Q(x)}$ where $P(x)$ and $Q(x)$ are polynomials (see Sec. 11)

is a *rational equation*. To solve it we must determine the roots of the algebraic equation $P(x)=0$ and choose among them those for which the denominator $Q(x)$ does not vanish.

An *irrational equation*, that is one involving x under a radical sign, is usually reducible to an algebraic equation by raising both sides of the equation to an appropriate power. This of course may lead to the appearance of extraneous roots, and therefore, after the roots of the resultant algebraic equation have been found, one should choose only those which satisfy the original equation.

III. *Transcendental equations*. An equation $f(x)=0$ is said to be *transcendental* if $f(x)$ is a transcendental function (see Sec. 11). Examples of such equations are logarithmic, exponential and trigonometric equations. Such an equation usually requires a special investigation to find out whether it possesses roots and what their number is.

76. **Testing a Root for Multiplicity.** Let us dwell in more detail on the properties of *multiple* roots (that is, having multiplicity higher than unity) of algebraic equations. Let a polynomial $P(x)$ possess a k -fold root α . Then factorization (**) in Sec. 75 can be written as

$$P(x) = (x - \alpha)^k Q(x) \text{ where } Q(\alpha) \neq 0$$

¹ Cardano, G. (1501-1576), an Italian scientist. Ferrari, L. (1522-1565), an Italian mathematician.

² The proof of this assertion was given by the famous Norwegian mathematician N.H. Abel (1802-1829) and the great French mathematician E. Galois (1811-1832).

The derivative $P'(x)$ (which is also a polynomial) is

$$P'(x) = (x - \alpha)^{k-1} [kQ(x) + (x - \alpha)Q'(x)] = (x - \alpha)^{k-1} Q_1(x)$$

where

$$Q_1(x) = kQ(x) + (x - \alpha)Q'(x)$$

Clearly, $Q_1(\alpha) = kQ(\alpha) \neq 0$, and therefore the derivative $P'(x)$ has the number α as its $(k-1)$ -fold root. If $k=1$ then $P'(\alpha) \neq 0$, and α is not a root of the derivative. Proceeding in this way we conclude that $P''(x)$ has α as its $(k-2)$ -fold root and so on. Accordingly, for the derivative $P^{(k-1)}(x)$ the number α is a simple root (of multiplicity 1), and for the k th derivative $P^{(k)}(x)$ this number is no longer a root.

Thus, if α is a root of multiplicity k for a polynomial $P(x)$ then

$$P(\alpha) = 0, P'(\alpha) = 0, \dots, P^{(k-1)}(\alpha) = 0 \text{ and } P^{(k)}(\alpha) \neq 0 \quad (*)$$

It can be proved that, conversely, if relations $(*)$ are fulfilled the number α is a k -fold root¹ of the polynomial $P(x)$.

Let the reader show that if α is a simple root of a polynomial equation $P(x) = 0$ the graph of the function $y = P(x)$ intersects the axis of abscissas without tangency and that if α is a two-fold root the graph touches the axis of abscissas and in a sufficiently small neighbourhood of the point of contact α it lies entirely on one side of the x -axis (this means that α is a point of extremum of the function $y = P(x)$). Similarly, for a three-fold root α the point $(\alpha, 0)$ of the axis of abscissas is a point of inflection of the graph, the tangent line at this point coinciding with the axis of abscissas. The case of a root α of multiplicity $k > 3$ can also be interpreted geometrically in like manner.

77. Approximate Solution of Equations. In this section we shall confine ourselves to some approximate procedures that are directly related to the content of this chapter, that is, are connected with the techniques applied to investigating the behaviour of functions. We begin with the remark that if there is some information about the behaviour of a function $y = f(x)$ this is usually sufficient for the determination of the intervals within which the graph of the function cuts the axis of abscissas, that is, in which the roots of the equation $f(x) = 0$ are located. The function $f(x)$ is supposed to be continuous, and therefore, if it has different signs at some points x_1 and x_2 it follows from the properties of continuous functions (see Sec. 35) that there is at least one zero of the function in the interval $[x_1, x_2]$.

¹ A k -fold root (a root of multiplicity k) is also termed a k -fold zero or a zero of multiplicity k .

Suppose that we have managed to determine an interval $[x_1, x_2]$ of that type which is so small that there is only one root of the equation $f(x)=0$ in it; we shall call such an interval $[x_1, x_2]$ an interval of isolation of the root. The fact that a chosen interval is an interval of isolation can usually be checked with the aid of the derivative $f'(x)$: if it retains constant sign the function $f(x)$ is monotone in that interval and its graph cuts the axis of abscissas only once.

Thus, we shall proceed from the assumption that a root x_0 of a given equation $f(x)=0$ has been in a certain way isolated within an interval $[x_1, x_2]$ ($x_1 < x_2$). Each of the numbers x_1 and x_2 can be regarded as an *approximation* to the root x_0 , namely, x_1 as a minor approximation and x_2 as a major approximation, the difference $x_2 - x_1$ serving as the limiting absolute error of these approximations. The approximate methods presented in this section are all based on a *reduction* of a given interval of isolation $[x_1, x_2]$ of a root x_0 of a given function $f(x)$ to a new interval of isolation $[x'_1, x'_2]$ such that

$$x_1 \leq x'_1 < x_0 < x'_2 \leq x_2$$

These techniques are thus based on the passage to a narrower interval of isolation. The new end points x'_1 and x'_2 are better approximations to the root x_0 than the former values x_1 and x_2 . The application to the interval $[x'_1, x'_2]$ of the same method or of another technique of the same kind leads to still better approximations x''_1 and x''_2 of the root x_0 , etc.

We shall discuss the three simple techniques known as the *cut-and-try method* (also spoken of as the *method of trials and errors*), the *method of chords* and the *method of tangents* (also known as *Newton's method*) and *combined methods* using combinations of these approaches.

It should also be noted that it is sometimes more convenient to rewrite the given equation $f(x)=0$ in an appropriately chosen form $\varphi_1(x)=\varphi_2(x)$ achieved by transposing some terms of the equation from left to right, which often facilitates the determination of an approximation to the sought-for root. If the graphs of the functions $\varphi_1(x)$ and $\varphi_2(x)$ are known the abscissas of their points of intersection are nothing but the roots of the original equation. For instance, on rewriting the equation $\ln x + x - 2 = 0$ in the form $\ln x = 2 - x$ and constructing the corresponding graphs we readily conclude that the original equation possess a single root contained between 1 and 2.

1. *Cut-and-try method*. This method is the simplest although not the best of the methods of constructing successive approximations to a root of an equation.

Let $[x_1, x_2]$ be an interval of isolation for a root x_0 of an equation $f(x)=0$. If the root is simple the values of the function at the end points of the interval must be of different signs; for definiteness, suppose that $f(x_1) < 0$ and $f(x_2) > 0$. Let us choose an arbitrary value $x=x'$ belonging to the interval $[x_1, x_2]$ and test it by substituting into the expression of the function $f(x)$; if $f(x') < 0$ we can replace x_1 by x' to obtain the narrower interval of isolation $[x', x_2]$; if $f(x') > 0$ we similarly arrive at the narrower interval of isolation $[x_1, x']$ by replacing x_2 by x' (it is unlikely that we obtain $f(x')=0$; in such a case we of course find the exact value of the root, but for our aims this rare case is of no interest).

Applying the cut-and-try method indefinitely we get an infinite sequence of points $x', x'', \dots$ which, as can be proved, has the root x_0 as its limit. This makes it possible to determine an approximate value of the root with an arbitrarily small error by applying the cut-and-try procedure an appropriate finite number of times.

Example. Let us take the equation

$$f(x) = x^3 + 1.1x^2 + 0.9x - 1.4 = 0$$

Since $f'(x) = 3x^2 + 2.2x + 0.9 > 0$ for all the values of x , the function $f(x)$ is monotone increasing, and hence its graph cuts Ox only once. Besides, $f(0) = -1.4$ and $f(1) = 1.6$, which means that there is a single real root located within the interval $[0, 1]$.

Let us compute $f(0.5) = -0.55$ and then $f(0.7) = 0.112$. This shows that $[0.5, 0.7]$ is a reduced interval of isolation of the sought-for root. Next we find

$$f(0.6) = -0.25, \quad f(0.65) = -0.076 \quad \text{and} \quad f(0.67) = -0.002$$

These approximations clearly approach the root; it lies in the interval $[0.67, 0.7]$. To improve the major approximation we can test the value $x=0.68$, which yields $f(0.68) = 0.034$. We have thus found the new interval of isolation $[0.67, 0.68]$ which is 100 times as short as the original interval $[0, 1]$. If the number 0.675 is chosen as an approximation of the root the absolute error is less than 0.005.

The cut-and-try method often involves more lengthy calculations than the methods of chords and tangents presented below since in the former method an improved value of the approximation is chosen, to a certain extent, *at random* whereas in the latter methods the choice is made in accordance with some *standard procedures*.

II. The method of chords. Suppose that the conditions of the problem are the same as in the cut-and-try method. Let us connect the ends of the arc $\widehat{M_1 M_2}$ (see Fig. 99) of the curve $y=f(x)$

corresponding to the interval $[x_1, x_2]$ by the chord M_1M_2 . In Fig. 99a it is clearly seen that the point of intersection $x = x'_1$ of that chord with Ox is closer to the point x_0 than to x_1 . Proceeding from the new (narrowed) interval $[x'_1, x_2]$ we obtain, in like manner, another point x''_1 which is still closer to x_0 than x'_1 . We thus receive a sequence of points $x_1, x'_1, x''_1, \dots$ which is increasing and tending to the

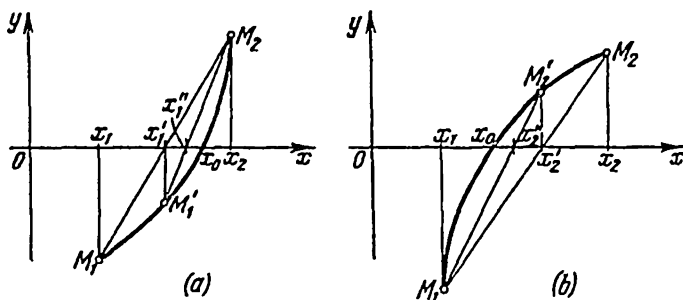


Fig. 99

sought-for root x_0 . In Fig. 99b we see an analogous sequence of points $x_2, x'_2, x''_2, \dots$ which is decreasing and tending to the root x_0 .

The equation of the chord M_1M_2 is written as

$$\frac{y - f(x_1)}{f(x_2) - f(x_1)} = \frac{x - x_1}{x_2 - x_1}$$

Putting $y=0$ we find the expression for the abscissa x' of the point of intersection of the chord with Ox :

$$x' = x_1 - f(x_1) \frac{x_2 - x_1}{f(x_2) - f(x_1)} \quad \text{or} \quad x' = x_2 - f(x_2) \frac{x_2 - x_1}{f(x_2) - f(x_1)} \quad (A)$$

Formula (A) is valid for both cases depicted in Fig. 99a and b (and also for the case $f(x_1) > 0, f(x_2) < 0$) and provides the new approximation x' to the root x_0 expressed in terms of the foregoing approximations x_1 and x_2 . To obtain a narrower interval of isolation it is sufficient to substitute x' for x_1 or x_2 . The question as to which of the points x_1 and x_2 should be replaced by x' can be decided by considering the behaviour of the function $f(x)$ (provided there is some information about it) or, when this is impossible, by examining the sign of $f(x')$.

Example. Let us apply the method of chords to the same equation

$$f(x) = x^3 + 1.1x^2 + 0.9x - 1.4 = 0$$

Putting $x_1 = 0$ and $x_2 = 1$ we find, by formula (A),

$$x^1 = 1 - f(1) \frac{1 - 0}{f(1) - f(0)} \approx 0.467$$

Next we put $x_1 = 0.467$ and $x_2 = 1$ and find

$$x^{II} \approx 0.617$$

Similarly,

$$x^{III} \approx 0.660, x^{IV} \approx 0.668, x^V \approx 0.670, x^{VI} \approx 0.670$$

The first three decimal digits in x^V and x^{VI} coinciding, we conclude (as almost always in calculations of this kind) that the approximations found closely approach the true value of the root. To estimate the accuracy let us test the value 0.671. We have $f(0.671) \approx 0.0012$, and, since $f(0.670) < 0$, the new interval of isolation of length 0.001 is $[0.670, 0.671]$. Choosing 0.6705 as an approximation to the root we obtain an error not exceeding 0.0005 which thus is 10 times smaller than that in the application of the cut-and-try method, the amount of calculations being approximately the same in both cases.

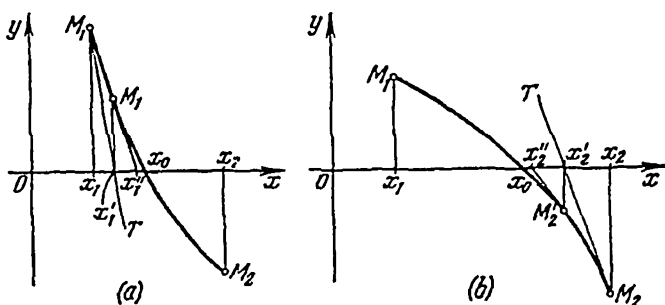


Fig. 100

III. The method of tangents. We shall suppose now that the arc $\widehat{M_1 M_2}$ of the curve $y = f(x)$ (see Fig. 100) corresponding to the interval of isolation $[x_1, x_2]$ possesses a tangent at its every point and has no points of inflection, that is, $f''(x)$ does not change sign within the interval $[x_1, x_2]$.

While the idea of the method of chords is to replace an arc by its chord, the method of tangents proceeds from the replacement of the arc by the corresponding segment of its tangent line. The tangent is drawn at the end point M_1 or M_2 of the arc, namely, at the one lying above Ox if the arc is concave (i.e. $f''(x) > 0$; see Fig. 100a) or below Ox if the arc is convex (i.e. $f''(x) < 0$; see Fig. 100b).

These conditions guarantee that the point of intersection x_1' (or x_2') of the tangent line with the x -axis is always between the root x_0 and one of the end points (x_1 or x_2) of the interval of isolation $[x_1, x_2]$. The new interval $[x_1', x_2]$ or $[x_1, x_2']$ serves as a reduced interval of isolation of the root x_0 . If one of the conditions we have posed is violated the new interval of isolation may happen

to be wider than the original one, and then we fail to approach the root closer than before. If these conditions hold the repeated application of the method results in the appearance of a sequence tending to x_0 as limit. Thus, this method as well makes it possible to determine the root with an arbitrary accuracy.

Fig. 100 demonstrates two possible cases; let the reader draw the corresponding sketches for the other two cases.

The equation of the tangent M_1T (or M_2T) is

$$y - f(x_1) = f'(x_1)(x - x_1) \quad (\text{or} \quad y - f(x_2) = f'(x_2)(x - x_2))$$

Putting $y=0$ we find the abscissa x'_1 (or x'_2) of the point of intersection of the tangent with Ox :

$$x'_1 = x_1 - \frac{f(x_1)}{f'(x_1)} \quad \left(\text{or} \quad x'_2 = x_2 - \frac{f(x_2)}{f'(x_2)} \right) \quad (B)$$

Example. Let us again take the equation

$$f(x) = x^3 + 1.1x^2 + 0.9x - 1.4 = 0$$

Since $f''(x) = 6x + 2.2 > 0$ in the interval $[0, 1]$, the tangent should be drawn at the point with abscissa equal to 1. Applying, in succession, formula (B), we get

$$x^I = 1 - \frac{f(1)}{f'(1)} \approx 0.738, \quad x^{II} = 0.738 - \frac{f(0.738)}{f'(0.738)} \approx 0.674, \\ x^{III} \approx 0.671, \quad x^{IV} \approx 0.671$$

The method of computing the approximations clearly indicates that $f(0.671) > 0$. Since $f(0.670) < 0$, the new interval of isolation is $[0.670, 0.671]$ which was obtained when we used the method of chords, but here the amount of calculations is still more reduced.

IV. Combined methods. The term "combined method" indicates that different methods of solving equations are used simultaneously.

We shall suppose that the conditions guaranteeing the possibility of the application of all the three methods discussed above are fulfilled: the arc of the curve $y=f(x)$ corresponding to a given interval of isolation $[x_1, x_2]$ of the root x_0 of the equation $f(x)=0$ has a tangent at its every point, has no points of inflection and $f(x_1)f(x_2) < 0$. If, for instance, we combine the methods of chords and tangents this results in two sequences of points tending to the point x_0 and providing minor and major approximations. In the case shown in Fig. 100a the method of tangents gives the minor approximations $x_1^{(m)}$ to the root x_0 whereas the method of chords the major approximations $x_2^{(m)}$; in the case depicted in Fig. 100b the situation is reverse. Such "two-sided" approximations accelerate the process of determining the root with a preassigned accuracy.

The combined method can be of course realized by using the method of chords or the method of tangents together with the cut-and-try method as was in fact done above in the solution of the equation

$$f(x) = x^3 + 1.1x^2 + 0.9x - 1.4 = 0$$

Let us apply to this equation the combination of the methods of chords and tangents. We have $x_1' \approx 0.467$ (see II) and $x_2' \approx 0.738$ (see III). Formulas (A) and (B) now yield

$$x_1'' = 0.467 - \frac{f(0.467)(0.738 - 0.467)}{f(0.738) - f(0.467)} \approx 0.658, \quad x_2'' \approx 0.674$$

and

$$x_1''' = 0.658 - \frac{f(0.658)(0.674 - 0.658)}{f(0.674) - f(0.658)} \approx 0.670, \quad x_2''' \approx 0.671$$

We see that this time the interval of isolation $[0.670, 0.671]$ whose length is 0.001 that of the original interval $[0, 1]$ is achieved at the third step.

In modern mathematics methods for approximate solution of equations are developed in one of its most important divisions termed *computational mathematics*. Besides the three methods and their combinations discussed in this section an extremely important role is played by the so-called *iteration method* (e.g. see [12]). (The methods of chords and tangents can be interpreted as special modifications of the iteration method.) Emphasis has been laid upon the development of techniques for solving algebraic equations providing not only real but also complex roots; one of the most frequently used methods of this kind was initiated by N. I. Lobachevsky. For these questions we refer the reader to [7], [12] and [13].

Modern electronic computers are successfully used to solve equations with high precision, and there are various standard programmes compiled for most frequently used approximate methods.

QUESTIONS

1. State and prove Fermat's theorem. Explain its geometrical interpretation.
2. State and prove Rolle's theorem. How is it interpreted geometrically?
3. State Lagrange's theorem; interpret it geometrically and give its analytical proof.
- 4*. State and prove Cauchy's theorem.
5. State and prove the direct and the converse theorems on the connection between increase or decrease of a function and the sign of its derivative. What is the geometrical meaning of these theorems?

6. State the definitions of points of extremum (points of maximum and minimum) of a function, of an extremal value of a function and of absolute extrema (the greatest and the least values).

7. What is the necessary condition for extremum? Give examples demonstrating that it is not sufficient.

8. What is the sufficient condition for extremum connected with the first derivative of a function?

9. Describe the general scheme for testing a function for extrema.

10. State the sufficient condition for extremum in terms of the second derivative and prove it.

11. How do we determine the greatest and the least values of a function in a given interval?

12. State the definitions of convexity and concavity of a function. What is a point of inflection?

13. State and prove the theorem on the connection between the character of the convexity of the graph of a function $y=f(x)$ and the sign of its second derivative $f''(x)$.

14. State the sufficient test for points of inflection.

15. Formulate L'Hospital's theorem. Give examples of its application to various limits.

16. What do we call an asymptote of a curve?

17. Derive analytical tests for determining vertical and inclined asymptotes of a curve $y=f(x)$.

18. Describe the general scheme for the investigation of the behaviour of a function.

19. Derive the formula for the differential of arc length. On what geometrical property of a curve is this formula based?

20. What is the curvature of a line at its given point? Derive the formula for the curvature.

21. Give the definitions of the radius, centre and circle of curvature.

22. State the definitions of the evolute and the evolute. Explain the techniques for their approximate construction.

23. Derive the equations of the tangent line and of the normal plane to a space curve.

24. State the definition and derive the equations of a cylindrical helix.

25. State the definitions of a vector function of a scalar argument and of its hodograph.

26. How do we define the limit of a vector function and its continuity? State the rule for determining its limit.

27. State the definition of the derivative of a vector function of scalar argument and explain the geometrical meaning of this notion. Derive the formula for the resolution of the derivative into components along the coordinate axes.

28*. What is the velocity and the acceleration of a point? Derive the formulas for the tangential and normal accelerations.

29. State the definitions of a complex function of a real argument, of its limit and continuity.

30. How is the derivative of a complex function defined? State the rule for determining the derivative.

31. State the definition of the exponential function with imaginary exponent. Describe its hodograph, prove its periodicity and derive the formula for its derivative.

32. What do we call the exponential form of a complex number?

33. Derive Euler's formulas expressing the trigonometric functions sine and cosine in terms of exponential functions.

34. Describe how a polynomial can be factorized into linear and quadratic factors.

35. State the definition of a multiple root of an algebraic equation. How can a given root be tested for multiplicity?

36. How are the intervals of isolation of the roots of an equation determined?

37. Describe the cut-and-try method, the method of chords, the method of tangents and the combined method for finding approximations to the value of a root of an equation.

Chapter V

INTEGRAL CALCULUS

§ 1. Indefinite Integral

78. Antiderivative. When differentiating a function we solve the problem of *finding the derivative of a given function*. Now we proceed to the inverse problem: *given the derivative of a function, to find this function*. As will be seen, the solution of the inverse problem is of great importance for mathematical analysis and its applications. We begin with the definition of the *antiderivative*.

Definition. An *antiderivative* (a *primitive*) of a given function $f(x)$ in a given interval is any function $F(x)$ whose derivative is the given function:

$$F'(x) = f(x)$$

Let us consider some examples elucidating the relationship between the given function and its antiderivative.

Let $y = x^2$. For which function does x^2 serve as its derivative? Obviously, for $\frac{x^3}{3}$:

$$\left(\frac{x^3}{3}\right)' = x^2$$

Therefore the function $\frac{x^3}{3}$ is an antiderivative of x^2 . However, the function x^2 possesses other antiderivatives. Indeed, the derivative of $\frac{x^3}{3} + 5$, of $\frac{x^3}{3} - 100$ and, generally, of $\frac{x^3}{3} + C$, where C is an arbitrary constant, is also equal to x^2 . Hence, any function of the form $\frac{x^3}{3} + C$ is an antiderivative of x^2 .

Thus, the connection between the functions x^2 and $\frac{x^3}{3} + C$, for any arbitrary constant C , is that the former is the *derivative* of the latter and the latter is an *antiderivative* of the former.

Similarly, for $x > 0$, any function of the form $\ln x + C$ is an antiderivative of $\frac{1}{x}$, any function of the form $\sin x + C$ is an antiderivative of $\cos x$, etc.

These examples indicate that a given function has infinitely many antiderivatives because the constant C can take on arbitrary values; on the other hand, a given function possesses, of course, only one derivative.

Now let us formulate the key theorem on antiderivatives.

Theorem. Every continuous function possesses an infinite number of antiderivatives, any two of them only differing by a constant.

The most difficult thing is to prove that a given continuous function has at least one antiderivative. This fact is not proved here, but later on we shall give its geometrical illustration. In what follows we assume that *any continuous function $f(x)$ has an antiderivative $F(x)$* . But then, for any constant C , the function $F(x) + C$ is also an antiderivative because

$$[F(x) + C]' = F'(x) = f(x)$$

Thus, every continuous function $f(x)$ has an infinitude of antiderivatives.

It remains to show that any two antiderivatives of the function $f(x)$ differ from each other only in a constant summand. Let $F(x)$ and $\Phi(x)$ be two antiderivatives of $f(x)$ which are not identically equal to each other. We have

$$F'(x) = f(x) \quad \text{and} \quad \Phi'(x) = f(x)$$

Subtracting the latter equality from the former we obtain $F'(x) - \Phi'(x) = 0$, i. e. $[F(x) - \Phi(x)]' = 0$. But if the derivative of a function (in our case of $F(x) - \Phi(x)$) is identically equal to zero, the function itself is constant (Sec. 59), and consequently $F(x) - \Phi(x) = C_1$, where C_1 is a concrete constant, which is what we set out to prove.

Thus, the expression $F(x) + C$ where $F(x)$ is a *concrete antiderivative* of $f(x)$ and C is an *arbitrary constant* embraces all the antiderivatives of $f(x)$. Making C assume different numerical values we obtain different antiderivatives.

Geometrical illustration of the existence of an antiderivative. Now we shall illustrate geometrically that any continuous function has an antiderivative.

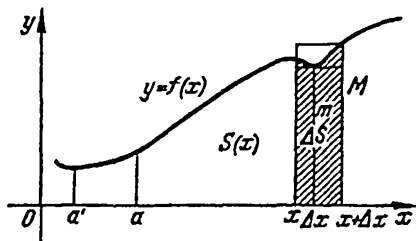


Fig. 101

Let us consider the graph of a continuous function $y=f(x)$ (see Fig. 101); we assume that $f(x) > 0$. Fix a point a and denote by $S(x)$ the area of the figure bounded by the segment $[a, x]$ of the axis of abscissas, by the graph of the function $y=f(x)$ and by two perpendiculars to the axis of abscissas at the points a and x . It is obvious that making x vary we change the area of the figure, which means that $S(x)$ is in fact a function of x . Let us find the derivative of the function $S(x)$. By the definition of the derivative, we have

$$S'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta S}{\Delta x}$$

Let x gain an increment Δx ; the area $S(x)$ then receives the corresponding increment ΔS (Fig. 101). Geometrically, it appears obvious that

$$m\Delta x < \Delta S < M\Delta x$$

where m and M are, respectively, the least and the greatest values of the function $f(x)$ in the interval $[x, x+\Delta x]$. For $m\Delta x$ is the area of the rectangle lying entirely inside the figure whose area is ΔS and $M\Delta x$ is the area of the rectangle enveloping this figure. Dividing all members of this inequality by Δx we obtain

$$m < \frac{\Delta S}{\Delta x} < M$$

The function $f(x)$ being continuous, both m and M (which are the least and the greatest values of the function in the interval $[x, x+\Delta x]$) tend to the value of the function at the point x , that is to $f(x)$, as $\Delta x \rightarrow 0$. By the test for the existence of a limit given in Sec. 30, the ratio $\frac{\Delta S}{\Delta x}$ tends to the same limit as $\Delta x \rightarrow 0$:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta S}{\Delta x} = f(x)$$

Thus, $S(x)$ is in fact an antiderivative of $f(x)$.

It can be shown in like manner that there exists an infinite set of antiderivatives. Taking as initial point not the point a but some other point a' we obtain a new function expressing the area of the figure with base $[a', x]$ which differs from $S(x)$ by the area of the figure with base $[a', a]$, that is by a constant summand.

This argument which is almost entirely based on geometrical considerations cannot of course serve as an analytic proof of the theorem since the notion of the area of an arbitrary figure was not strictly defined beforehand. At the same time it shows that the operation of finding antiderivatives or, as we shall call it, *integration* is closely related to the problem of computing areas bounded by curves.

79. Indefinite Integral. Basic Table of Integrals.

Definition. The operation of finding antiderivatives is called *indefinite integration*, and the collection of all antiderivatives of a given function $f(x)$ is called the *indefinite integral* of $f(x)$ and is denoted by the symbol

$$\int f(x) dx$$

The function $f(x)$ is called the *integrand*, the expression $f(x) dx$ is the *element of integration* and the variable x is the *variable of integration*.

It is supposed that the integrand $f(x)$ is continuous for the values of x under consideration.

According to what was shown in Sec. 78, we have

$$\int f(x) dx = F(x) + C$$

where $F(x)$ is an arbitrary (concrete) antiderivative of $f(x)$ and C is an arbitrary constant.

Indefinite integration is the inverse of the operation of differentiation. By means of differentiation we find the derivative of a given function and by means of indefinite integration we find the antiderivatives corresponding to a given derivative. The correctness of integration can always be checked by differentiating the result. The meaning of the symbol $\int f(x) dx$ in the notation of the result of the operation inverse to differentiation will be explained later on.

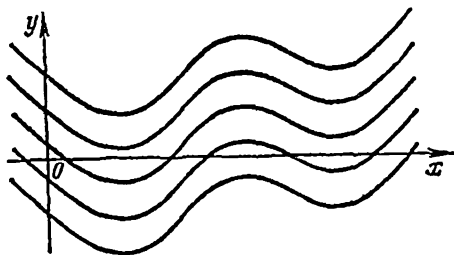


Fig. 102

To find the indefinite integral of a function means to find all its antiderivatives (for which it is sufficient to determine one of them). This is the reason why we speak of *indefinite integration* as it is not indicated which of the antiderivatives is meant.

The graph of an antiderivative of a function $f(x)$ is called an *integral curve* of the function $y=f(x)$. It is obvious that any integral curve can be obtained by a translation (parallel displacement) of any other integral curve in the vertical direction. Therefore, geometrically, the indefinite integral is represented by the set of all integral curves obtained by continuous parallel displacement of one of them in the vertical direction (Fig. 102).

By the definition of the indefinite integral we have

$$\left(\int f(x) dx \right)' = f(x) \quad \text{or, equivalently,} \quad d \int f(x) dx = f(x) dx$$

and

$$\int f'(x) dx = f(x) + C \quad \text{or} \quad \int df(x) = f(x) + C$$

Hence, to within the constant of integration entering into the indefinite integral, when written consecutively, the symbols of the differential and of the indefinite integral cancel out. Throughout this section the terms "integral" and "integration" are understood in the sense of an "indefinite integral" and "indefinite integration", the integrand supposed to be continuous in the given interval.

First of all we shall write down the integration formulas directly implied by the differentiation formulas for the basic elementary functions. Each of them is easily checked by differentiation.

The Basic Table of Integrals.

$$(1) \int x^k dx = \frac{x^{k+1}}{k+1} + C, \quad k \neq -1.$$

$$(2) \int \frac{dx}{x} = \ln|x| + C.^1$$

$$(3) \int a^x dx = \frac{a^x}{\ln a} + C.$$

$$(4) \int e^x dx = e^x + C.$$

$$(5) \int \cos x dx = \sin x + C.$$

$$(6) \int \sin x dx = -\cos x + C.$$

$$(7) \int \frac{dx}{\cos^2 x} = \tan x + C.$$

$$(8) \int \frac{dx}{\sin^2 x} = -\cot x + C.$$

$$(9) \int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C.$$

$$(10) \int \frac{dx}{1+x^2} = \arctan x + C.$$

$$(11) \int \frac{dx}{\sqrt{x^2 \pm 1}} = \ln(x + \sqrt{x^2 \pm 1}) + C.$$

$$(12) \int \frac{dx}{x^2-1} = \frac{1}{2} \ln \left| \frac{x-1}{x+1} \right| + C.$$

¹ In this formula x can be in the interval $(-\infty, 0)$ or $(0, \infty)$ but cannot vary in an interval containing the point 0 because in such an interval the integrand is no longer continuous. If $x > 0$ we have $\ln|x| = \ln x$, and the derivative is $\frac{1}{x}$. For $x < 0$ we have $\ln|x| = \ln(-x)$, and the derivative is again equal to $\frac{1}{x}$. Thus, the formula embraces both cases.

Below are some more integration formulas connected with hyperbolic and inverse hyperbolic functions (see Sec. 23):

$$\begin{aligned}\int \cosh x \, dx &= \sinh x + C, \quad \int \frac{dx}{\cosh^2 x} = \tanh x + C, \\ \int \sinh x \, dx &= \cosh x + C, \quad \int \frac{dx}{\cosh^2 x} = -\coth x + C, \\ \int \frac{dx}{\sqrt{x^2+1}} &= \sinh^{-1} x + C = \ln(x + \sqrt{x^2+1}) + C, \\ \int \frac{dx}{\sqrt{x^2-1}} &= \cosh^{-1} x + C = \ln(x + \sqrt{x^2-1}) + C\end{aligned}$$

(The last two formulas lead to Formula (11) of the basic table.)

$$\int \frac{dx}{1-x^2} = \tanh^{-1} x + C = \frac{1}{2} \ln \frac{1+x}{1-x} + C$$

The latter formula is valid for the interval $-1 < x < 1$ which is the domain of definition of $\tanh^{-1} x$; Formula (12) is its modification and generalization.

80. Simplest Integration Rules.

Theorem I (On Integrating a Sum). The integral of a sum of a finite number of functions is equal to the sum of the integrals of these functions:

$$\int (u + v + \dots + w) \, dx = \int u \, dx + \int v \, dx + \dots + \int w \, dx \quad (*)$$

where $u, v, \dots, w$ are functions of the independent variable x .

An equality whose both members are indefinite integrals means that its left and right members are the collections of all the antiderivatives of one and the same function. Consequently, to check such an equality it is sufficient to verify that the derivatives of its left and right members are equal to each other. Therefore, applying to the right member of equality (*) the rule for differentiating a sum we obtain

$$\begin{aligned}& \left[\int u \, dx + \int v \, dx + \dots + \int w \, dx \right]' = \\ &= \left(\int u \, dx \right)' + \left(\int v \, dx \right)' + \dots + \left(\int w \, dx \right)' = u + v + \dots + w\end{aligned}$$

which coincides with the integrand of the integral on the left-hand side of the equality, i. e. with its derivative. The theorem has been proved.

Theorem II (On Taking a Constant Factor Outside the Integral Sign). A constant factor in the integrand can be taken outside the sign of indefinite integration:

$$\int c f(x) \, dx = c \int f(x) \, dx \quad (c = \text{const})$$

Indeed, the derivative of the right-hand side of the equality, i.e.

$$\left(c \int f(x) dx\right)' = c \left(\int f(x) dx\right)' = cf(x)$$

is equal to the derivative of the left-hand side, which is what we set out to prove.

Example.

$$\begin{aligned} \int (2 \sin x - 3 \cos x) dx &= 2 \int \sin x dx - 3 \int \cos x dx = \\ &= -2 \cos x - 3 \sin x + C \end{aligned}$$

Though every intermediate integration contributes an arbitrary constant term to the final result we write only one such term since the sum of arbitrary constants is also an arbitrary constant.

Theorem III (On the Invariance of Integration Formulas). Let

$$\int f(x) dx = F(x) + C$$

be any given integration formula and $u = \varphi(x)$ any function possessing a continuous derivative. Then¹

$$\int f(u) du = F(u) + C$$

Proof. From $\int f(x) dx = F(x) + C$ it follows that $F'(x) = f(x)$. Now take the function $F(u) = F[\varphi(x)]$; by the theorem on the invariance of the form of the first differential of a function (Sec. 51, III), its differential is

$$dF(u) = F'(u) du = f(u) du$$

Hence

$$\int f(u) du = \int dF(u) = F(u) + C$$

This rule is extremely important. It shows that the basic table of integrals is valid irrespective of whether the variable of integration is an independent variable or any function of another variable possessing a continuous derivative with respect to that variable and thus it extends considerably the range of application of the table.

For instance, take the integral $\int 2xe^{x^2} dx$. Observing that $2x dx$ is the differential $d(x^2)$ we rewrite the integral as

$$\int 2xe^{x^2} dx = \int e^{x^2} d(x^2) = \int e^u du$$

¹ More precisely, it must be additionally required that the values of the function $u = \varphi(x)$ should lie in the interval where the function f is defined and continuous.

where $u = x^2$. By Formula (4) of the basic table and Theorem III, the latter integral is equal to $e^u + C$, and hence

$$\int 2xe^{x^2} dx = e^{x^2} + C$$

This example shows that it is advisable to transform a given element of integration to an element of integration of a simpler integral, for instance, to one included into the basic table.

It should be stressed that integration is a much more complicated operation than differentiation. Elementary functions are differentiated according to definite rules and formulas while integration involves, so to say, an "individual" approach to every function. It requires much experience to find an appropriate transformation of the element of integration reducing the given integral to a tabular one¹. Here we give a number of examples intended to help the reader to acquire such experience. After studying these examples the reader should solve as many simple problems of this kind as possible before proceeding to the general methods of integration presented in the following sections.

Examples. (1) Let us find $\int \sin 5x dx$. On multiplying and dividing the integrand by 5 and taking the factor $\frac{1}{5}$ outside the integral symbol we obtain

$$\int \sin 5x dx = \frac{1}{5} \int \sin 5x \cdot 5 dx = \frac{1}{5} \int \sin 5x d(5x)$$

Now, putting $5x = u$ we arrive at an integral included into the basic table:

$$\int \sin 5x dx = \frac{1}{5} \int \sin u du = -\frac{1}{5} \cos u + C = -\frac{1}{5} \cos 5x + C$$

We similarly find

$$\begin{aligned} \int e^{-3x} dx &= -\frac{1}{3} \int e^{-3x} d(-3x) = -\frac{1}{3} \int e^u du = \\ &= -\frac{1}{3} e^u + C = -\frac{1}{3} e^{-3x} + C \end{aligned}$$

(2) Consider the integral $\int (2x-1)^{100} dx$. On multiplying and dividing it by 2 and taking into account that $2dx = d(2x-1)$ we

¹ It should be noted that such a transformation is by far not always possible. This is connected with a specific difficulty of the integral calculus: an antiderivative of a comparatively simple function may be inexpressible in terms of elementary functions, that is, it may be a function of a more complex type. This question will be discussed in more detail in Sec. 85.

receive

$$\int (2x-1)^{100} dx = \frac{1}{2} \int (2x-1)^{100} d(2x-1)$$

The change $2x-1=u$ now leads to

$$\int (2x-1)^{100} dx = \frac{1}{2} \int u^{100} du = \frac{1}{2} \frac{u^{101}}{101} + C = \frac{1}{202} (2x-1)^{101} + C$$

The advantage of the technique applied is readily seen if we compare this procedure with the integration of the polynomial of the 100th degree appearing from Newton's binomial formula for $(2x-1)^{100}$.

(3) Let us find the integral $\int \frac{x}{\sqrt{x^2+1}} dx$. To this end, we represent the integral in the form

$$\frac{1}{2} \int \frac{2x}{\sqrt{x^2+1}} dx = \frac{1}{2} \int (x^2+1)^{-\frac{1}{2}} d(x^2+1)$$

Putting (mentally) $x^2+1=u$ we obtain

$$\int \frac{x}{\sqrt{x^2+1}} dx = \frac{1}{2} \frac{(x^2+1)^{\frac{1}{2}}}{\frac{1}{2}} + C = \sqrt{x^2+1} + C$$

After some training it becomes unnecessary to write down all the intermediate calculations whose greater part can be done mentally.

(4) For the integral $\int \frac{dx}{3x-1}$ we obtain

$$\int \frac{dx}{3x-1} = \frac{1}{3} \int \frac{(3x-1)'}{3x-1} dx = \frac{1}{3} \int \frac{d(3x-1)}{3x-1} = \frac{1}{3} \ln |3x-1| + C$$

Generally, if the integrand is a fraction whose numerator is the derivative of the denominator, the integral is equal to the logarithm of the absolute value of the denominator.

Indeed,

$$\int \frac{f'(x)}{f(x)} dx = \int \frac{df(x)}{f(x)} = \int \frac{du}{u} = \ln |u| + C = \ln |f(x)| + C$$

For instance,

$$\int \frac{x}{1+x^2} dx = \frac{1}{2} \int \frac{2x dx}{1+x^2} = \frac{1}{2} \ln (1+x^2) + C$$

$$\int \tan x dx = \int \frac{\sin x}{\cos x} dx = -\ln |\cos x| + C$$

$$\int \cot x dx = \int \frac{\cos x}{\sin x} dx = \ln |\sin x| + C$$

$$\int \frac{e^x - e^{-x}}{e^x + e^{-x}} dx = \ln (e^x + e^{-x}) + C$$

(5) Taking the integral $\int \frac{x}{1+x^4} dx$ and noting that $x dx = \frac{1}{2} d(x^2)$ we receive

$$\int \frac{x dx}{1+x^4} = \frac{1}{2} \int \frac{d(x^2)}{1+(x^2)^2} = \frac{1}{2} \arctan(x^2) + C$$

(6) On performing the division of the numerator by the denominator in the integral $\int \frac{3x+2}{2x-1} dx$ we obtain 1.5 in the quotient and 3.5 in the remainder. Consequently,

$$\int \frac{3x+2}{2x-1} dx = \int \left(1.5 + \frac{3.5}{2x-1} \right) dx = 1.5x + 1.75 \ln|2x-1| + C$$

(7) Let us find the integrals $\int \frac{dx}{x^2-a^2}$ and $\int \frac{dx}{x^2+a^2}$. We have

$$\begin{aligned} \int \frac{dx}{x^2-a^2} &= \frac{1}{a^2} \int \frac{dx}{\left(\frac{x}{a}\right)^2-1} = \frac{1}{a} \int \frac{d\left(\frac{x}{a}\right)}{\left(\frac{x}{a}\right)^2-1} = \\ &= \frac{1}{a} \frac{1}{2} \ln \left| \frac{\frac{x}{a}-1}{\frac{x}{a}+1} \right| + C = \frac{1}{2a} \ln \left| \frac{x-a}{x+a} \right| + C \end{aligned}$$

The second integral is found similarly:

$$\int \frac{dx}{x^2+a^2} = \frac{1}{a} \arctan \frac{x}{a} + C$$

An integral of the type $\int \frac{dx}{x^2+px+q}$ can readily be reduced to these integrals; to this end, the denominator should be represented as a sum or a difference of squares. For example:

$$\int \frac{dx}{x^2+4x+8} = \int \frac{dx}{(x+2)^2+4} = \frac{1}{2} \arctan \frac{x+2}{2} + C$$

and

$$\int \frac{dx}{x^2+4x+1} = \int \frac{dx}{(x+2)^2-3} = \frac{1}{2\sqrt{3}} \ln \left| \frac{x+2-\sqrt{3}}{x+2+\sqrt{3}} \right| + C$$

(8) Consider $\int \frac{x+3}{x^2+4x+8} dx$. Observing that the derivative of the denominator is $2x+4$ we write the numerator in the form $\frac{1}{2}(2x+4)+1$ and split the given integral into two integrals reducible to tabular ones as was shown above:

$$\begin{aligned} \int \frac{x+3}{x^2+4x+8} dx &= \frac{1}{2} \int \frac{2x+4}{x^2+4x+8} dx + \int \frac{dx}{x^2+4x+8} = \\ &= \frac{1}{2} \ln(x^2+4x+8) + \frac{1}{2} \arctan \frac{x+2}{2} + C \end{aligned}$$

(9) The integrals $\int \frac{dx}{\sqrt{a^2 - x^2}}$ and $\int \frac{dx}{\sqrt{x^2 \pm a^2}}$ are found by analog, with Example 7:

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \frac{1}{a} \int \frac{dx}{\sqrt{1 - \left(\frac{x}{a}\right)^2}} = \int \frac{d\left(\frac{x}{a}\right)}{\sqrt{1 - \left(\frac{x}{a}\right)^2}} \arcsin \frac{x}{a} + C$$

and

$$\begin{aligned} \int \frac{dx}{\sqrt{x^2 \pm a^2}} &= \ln \left(\frac{x}{a} + \sqrt{\left(\frac{x}{a}\right)^2 \pm 1} \right) + C = \\ &= \ln (x + \sqrt{x^2 \pm a^2}) + C \end{aligned}$$

(When transforming the expression involving the logarithm we include the term $-\ln a$ into the arbitrary constant.)

(10) The integral $\int \sin x \cos x dx$ is found with the aid of the well-known trigonometric formula:

$$\int \sin x \cos x dx = \frac{1}{2} \int \sin 2x dx = \frac{1}{4} \int \sin 2x d(2x) = -\frac{1}{4} \cos 2x + C$$

For this integral we can use another method:

$$\int \sin x \cos x dx = \int \sin x d(\sin x) = \frac{1}{2} \sin^2 x + C$$

or

$$\int \sin x \cos x dx = -\int \cos x d(\cos x) = -\frac{1}{2} \cos^2 x + C$$

It may seem that for one and the same integral we have obtained three different expressions

$$-\frac{1}{4} \cos 2x + C, \quad \frac{1}{2} \sin^2 x + C \quad \text{and} \quad -\frac{1}{2} \cos^2 x + C$$

However, it can be easily checked that the difference between any two of them is constant. Therefore each of these expressions describes the set of all antiderivatives of $\sin x \cos x$.

(11) The integrals $\int \sin^2 x dx$ and $\int \cos^2 x dx$ are readily found with the help of the formulas $\sin^2 x = \frac{1 - \cos 2x}{2}$ and $\cos^2 x = \frac{1 + \cos 2x}{2}$. For example:

$$\int \cos^2 x dx = \frac{1}{2} \int (1 + \cos 2x) dx = \frac{1}{2} \left(x + \frac{\sin 2x}{2} \right) + C$$

(12) For the integral $\int \cos^2 x dx$ we find, in succession,

$$\begin{aligned}\int \cos^2 x dx &= \int \cos^2 x d(\sin x) = \int (1 - \sin^2 x) d(\sin x) = \\ &= \int d(\sin x) - \int \sin^2 x d(\sin x) = \sin x - \frac{1}{3} \sin^3 x + C\end{aligned}$$

81. Integration by Parts and by Change of Variable. Since the operation of integration is inverse to differentiation, each differentiation rule generates the corresponding integration rule. The three integration rules proved in Sec. 80 correspond to the rules for differentiating a sum of functions, a product of a function by a constant and a composite function.

I. Integration by parts. The method of integration by parts is implied by the formula for differentiating a product of two functions:

Let $u(x)$ and $v(x)$ be functions of x possessing continuous derivatives. We have

$$d(uv) = u dv + v du$$

whence

$$u dv = d(uv) - v du$$

Integrating both sides of the latter equality, we receive

$$\int u dv = \int d(uv) - \int v du$$

that is,

$$\int u dv = uv - \int v du \quad (*)$$

This is the formula of integration by parts which can be written at length as

$$\int u(x) v'(x) dx = u(x)v(x) - \int v(x) u'(x) dx$$

We do not write down the arbitrary constant appearing in the integration of $d(uv)$ and include it into the arbitrary constant in the second integral on the right-hand side of equality (*).

Integration by parts reduces to an appropriate representation of the element of integration $f(x)dx$ as a product of two factors $u(x)$ and $v'(x)dx = dv$ and to two integrations: (1) the determination of v from the expression for dv and (2) the determination of the integral of $v du$. It may turn out that these two integrations are easily carried out whereas it is difficult to find the original integral directly.

Examples. (1) Consider the integral $\int x e^x dx$. Let us take $e^x dx$ as dv and x as u ; we thus put:

$$e^x dx = dv, \quad x = u$$

whence

$$v = \int e^x dx = e^x \quad \text{and} \quad du = dx$$

By formula (*) we obtain

$$\int x e^x dx = x e^x - \int e^x dx = x e^x - e^x + C = e^x (x - 1) + C$$

The given integral has thus been reduced to tabular ones. It would be impractical to represent the element of integration as $x dx = dv$, $e^x = u$ since this would lead to $v = \frac{x^2}{2}$, $du = e^x dx$ and

$$\int x e^x dx = \frac{x^2}{2} e^x - \frac{1}{2} \int x^2 e^x dx$$

The resultant integral on the right-hand side of the last equality is even more complex than the original one. It is obvious that the latter integral, in its turn, can be reduced, by an appropriate application of this method, to the simpler integral on the left-hand side which has been found above (to this end we put $dv = e^x dx$ and $u = x^2$).

(2) Let us find $\int x \ln x dx$. We put

$$x dx = dv \quad \text{and} \quad \ln x = u$$

whence

$$v = \frac{x^2}{2} \quad \text{and} \quad du = \frac{dx}{x}$$

Then

$$\int x \ln x dx = \frac{x^2}{2} \ln x - \frac{1}{2} \int x dx = \frac{x^2}{2} \ln x - \frac{x^2}{4} + C$$

Here the desired result would not be obtained if we put $\ln x dx = dv$, $x = u$ because the integral $\int \ln x dx$ is not simpler than the original one; it must itself be found with the help of integration by parts by putting $dx = dv$, $\ln x = u$. Then $v = x$, $du = \frac{dx}{x}$ and $\int \ln x dx = x \ln x - x + C$.

Sometimes to obtain the desired result it is necessary to use integration by parts several times in succession.

(3) Take the integral $\int x^2 \cos x dx$. Here we put

$$\cos x dx = dv, \quad x^2 = u$$

and

$$v = \sin x, \quad du = 2x dx$$

Formula (*) yields

$$\int x^2 \cos x dx = x^2 \sin x - 2 \int x \sin x dx$$

The last integral is again found with the aid of integration by parts. Finally we receive

$$\int x^2 \cos x \, dx = x^2 \sin x + 2x \cos x - 2 \sin x + C$$

In practical work it is unnecessary to put down all the intermediate steps.

Repeated integration by parts makes it possible to find the integrals

$$\int x^m \sin x \, dx, \quad \int x^m \cos x \, dx, \quad \int x^m e^x \, dx$$

where m is a positive integer, and hence the integrals

$$\int P(x) \sin x \, dx, \quad \int P(x) \cos x \, dx, \quad \int P(x) e^x \, dx$$

where $P(x)$ is any polynomial as well.

Repeated integration by parts sometimes leads to the original integral, and then we either arrive at a useless identity (if the repeated integration has been done improperly) or at an equality that enables us to find the sought-for integral. Let us consider an example demonstrating this technique.

(4) We take the integral $\int e^x \cos x \, dx$ and put

$$e^x \, dx = dv, \quad \cos x = u$$

whence

$$v = e^x, \quad du = -\sin x \, dx$$

It follows that

$$\int e^x \cos x \, dx = e^x \cos x + \int e^x \sin x \, dx$$

Let us again apply integration by parts. On putting¹

$$e^x \, dx = dv, \quad \sin x = u,$$

and

$$v = e^x, \quad du = \cos x \, dx$$

we obtain

$$\int e^x \sin x \, dx = e^x \sin x - \int e^x \cos x \, dx$$

and arrive at the original integral as a summand on the right-hand side. Substituting the expression thus obtained into the result of the first integration we receive

$$\int e^x \cos x \, dx = e^x \cos x + e^x \sin x - \int e^x \cos x \, dx$$

¹ If we take $\sin x \, dx$ as dv here, the result will be a trivial identity

$$\int e^x \cos x \, dx = \int e^x \cos x \, dx$$

On transposing the integral from the right-hand side to the left we ultimately find

$$\int e^x \cos x \, dx = \frac{1}{2} e^x (\cos x + \sin x) + C$$

Experience enables one to see in which cases and in what way integration by parts should be applied.

II. *Integration by change of variable (by substitution).* When computing integrals we have already resorted to theorem III of Sec. 80 on the invariance of integration formulas. If we succeed in writing the element of integration in the form

$$f[\varphi(x)] \varphi'(x) \, dx = f(u) \, du$$

where $u = \varphi(x)$ and if the integral

$$\int f(u) \, du = F(u) + C$$

of the expression on the right-hand side is known, the original integral is equal to

$$\int f[\varphi(x)] \varphi'(x) \, dx = F[\varphi(x)] + C$$

The main difficulty in the application of this technique lies in the transformation of the element of integration because the desired result and the function $\varphi(x)$ are not known in advance. We considered various examples of such transformations in sec. 80 where in simple cases the change of variable $u = \varphi(x)$ was not written down and the final expression for the integral was found directly.

In more complicated cases it is advisable first to choose a substitution which seems to be applicable and then, after completing the transformation of the element of integration, to check whether we have succeeded in simplifying the integral. It may turn out that the new integral is not yet a tabular one, but if it can be reduced to such more easily than the original, the substitution is justified.

For instance, the integral $\int \frac{\sin 2x}{1 + \sin^4 x} \, dx$ is reduced by the substitution $\sin x = u$, $\cos x \, dx = du$ to the integral $\int \frac{2u}{1 + u^4} \, du$. On writing the latter in the form $\int \frac{d(u^2)}{1 + (u^2)^2}$ (the substitution $u^2 = v$ is done mentally) we see that it is equal to $\arctan u^2 + C$, and consequently,

$$\int \frac{\sin 2x}{1 + \sin^4 x} \, dx = \arctan (\sin^2 x) + C$$

(The more complicated substitution $\sin^2 x = u$ would bring the given integral to a tabular one immediately.)

Up to now we have used the method of substitution by replacing the variable of integration x by another variable u with the aid of a formula $u = \varphi(x)$. But it is also possible to make a substitution not by expressing u in terms of x but, by taking x as a function of u , that is, by putting

$$x = \psi(u), \quad dx = \psi'(u) du$$

(it is supposed that $\psi(u)$ and $\psi'(u)$ are continuous). Then

$$f(x) dx = f[\psi(u)] \psi'(u) du$$

and

$$\int f(x) dx = \int f[\psi(u)] \psi'(u) du \quad (**)$$

If the integral on the right-hand side of (**) is found and expressed as $F(u) + C$, the given integral can be found by returning to the variable x . To this end it is necessary to express u in terms of x from the equation $x = \psi(u)$, i.e. to determine the inverse function¹ and substitute it for u in the expression of the integral on the right-hand side of (**).

As before, the most difficult thing is to choose an appropriate substitution $x = \psi(u)$. Here are some examples illustrating both types of change of variable.

Examples. (1) Consider the integral $\int x^2 \sqrt[3]{4-3x^3} dx$. Let us put $4-3x^3 = u$. On differentiating we obtain $-9x^2 dx = du$ and $x^2 dx = -\frac{1}{9} du$. Therefore,

$$\begin{aligned} \int x^2 \sqrt[3]{4-3x^3} dx &= -\frac{1}{9} \int \sqrt[3]{u} du = -\frac{1}{12} u^{\frac{4}{3}} + C = \\ &= -\frac{1}{12} (4-3x^3)^{\frac{4}{3}} \sqrt[3]{4-3x^3} + C \end{aligned}$$

(2) For the integral $\int \frac{dx}{\sqrt{e^x+1}}$ we put $\sqrt{e^x+1} = u$, i.e. $e^x + 1 = u^2$. Then

$$e^x dx = 2u du \quad \text{and} \quad dx = \frac{2u du}{e^x} = \frac{2u du}{u^2-1}$$

Consequently,

$$\int \frac{dx}{\sqrt{e^x+1}} = \int \frac{2u du}{(u^2-1)u} = 2 \int \frac{du}{u^2-1} = \ln \left| \frac{u-1}{u+1} \right| + C$$

¹ The function $\psi(u)$ is usually assumed to be monotone.

On returning to the variable x we finally obtain

$$\int \frac{dx}{\sqrt{e^x+1}} = \ln \left| \frac{\sqrt{e^x+1}-1}{\sqrt{e^x+1}+1} \right| + C$$

Sometimes together with the change of variable we should apply some other methods (for instance, integration by parts) to complete the solution of the problem.

(3) Let us find the integral $\int x^5 e^{x^3} dx$. Here we put $x^3 = u$ and obtain $3x^2 dx = du$ whence

$$\int x^5 e^{x^3} dx = \frac{1}{3} \int u e^u du$$

The latter integral is already known (see I, Example (1)) where it is found with the aid of integration by parts), and we obtain

$$\int x^5 e^{x^3} dx = \frac{1}{3} e^u (u-1) + C = \frac{1}{3} e^{x^3} (x^3-1) + C$$

Up to now we have used substitutions of the type $u = \varphi(x)$. Now we shall illustrate the application of substitutions of the type $x = \psi(u)$ by examples of *trigonometric* and *hyperbolic substitutions* which are frequently used.

(4) Let us find $\int \sqrt{1-x^2} dx$. Putting $x = \sin u$, $dx = \cos u du$ and making the substitution we find

$$\int \sqrt{1-x^2} dx = \int \cos^2 u du = \frac{1}{2} \left(u + \frac{1}{2} \sin 2u \right) + C$$

(see Example 11, Sec. 80). Returning to the variable x and noting that $\sin 2u = 2 \sin u \cos u = 2x\sqrt{1-x^2}$ we obtain

$$\int \sqrt{1-x^2} dx = \frac{1}{2} (\arcsin x + x\sqrt{1-x^2}) + C$$

It should be noted that here the function $x = \sin u$ is only considered on the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ where it is monotone. Then x varies within the interval $[-1, 1]$, i.e. within the interval where the integrand is defined.

(5) Let us consider the integral $\int \sqrt{x^2+1} dx$. Here it is convenient to apply the substitution¹ $x = \sinh u$, $dx = \cosh u du$:

$$\begin{aligned} \int \sqrt{x^2+1} dx &= \int \cosh^2 u du = \frac{1}{2} \int (1 + \cosh 2u) du \\ &= \frac{1}{2} \left(u + \frac{1}{2} \sinh 2u \right) + C \end{aligned}$$

¹ Another possible substitution $x = \tan u$ leads to very complicated transformations. The introduction of hyperbolic functions substantially simplifies the transformations.

Since $u = \sinh^{-1} x = \ln(x + \sqrt{x^2 + 1})$ (see Sec. 23, II), we have

$$\sinh 2u = 2 \sinh u \cosh u = 2x \sqrt{x^2 + 1} \quad \text{and} \\ \int \sqrt{x^2 + 1} dx = \frac{1}{2} [x \sqrt{x^2 + 1} + \ln(x + \sqrt{x^2 + 1})] + C$$

Similarly, by means of the substitution $x = \cosh u$ we find

$$\int \sqrt{x^2 - 1} dx = \frac{1}{2} [x \sqrt{x^2 - 1} - \ln(x + \sqrt{x^2 - 1})] + C$$

The integrals in Examples 4 and 5 can be computed with the aid of integration by parts. Let us take the integral $\int \sqrt{x^2 + A} dx$ and put $\sqrt{x^2 + A} = u$, $dx = dv$; then $du = \frac{x}{\sqrt{x^2 + A}} dx$, $v = x$ and

$$\int \sqrt{x^2 + A} dx = x \sqrt{x^2 + A} - \int \frac{x^2}{\sqrt{x^2 + A}} dx$$

Now let us transform the integral on the right-hand side in the following way:

$$\int \frac{x^2}{\sqrt{x^2 + A}} dx = \int \frac{x^2 + A - A}{\sqrt{x^2 + A}} dx = \int \sqrt{x^2 + A} dx - A \ln(x + \sqrt{x^2 + A})$$

(see Example 9 in Sec. 80). Substituting this result into the preceding equality we arrive at an equation for the sought-for integral and find from it the expression

$$\int \sqrt{x^2 + A} dx = \frac{1}{2} [x \sqrt{x^2 + A} + A \ln(x + \sqrt{x^2 + A})] + C$$

Generally, the integration process reduces to transforming the given integral to an integral already known by means of an *appropriate application* of the techniques discussed above (an algebraic transformation of an element of integration, integration by parts and integration by the change of variable).

82. Integration of Rational Functions. One of the most important classes of elementary functions whose integrals are found by means of a definite and comparatively simple sequence of standard operations is the class of *rational functions* (*rational fractions*; see Sec. 11). Every rational function $R(x)$ has the form

$$R(x) = \frac{P(x)}{Q(x)}$$

where $P(x)$ and $Q(x)$ are *polynomials*.

First of all it should be noted that if the degree m of the numerator $P(x)$ is greater than or equal to the degree n of the denominator $Q(x)$, the division of the polynomial $P(x)$ by the

polynomial $Q(x)$ results in a polynomial $N(x)$ in the quotient and in a polynomial $P_1(x)$ of degree not higher than $(n-1)$ in the remainder. Consequently,

$$\frac{P(x)}{Q(x)} = N(x) + \frac{P_1(x)}{Q(x)}$$

The integration of the polynomial $N(x)$ does not present any difficulties, and hence the problem reduces to integrating a rational function in which the degree of the numerator is less than that of the denominator. Thus, we can limit ourselves to the case when the degree of the numerator is less than the degree of the denominator, that is, $m < n$. Such a rational function is called a *proper rational fraction*.

I. Decomposition of a proper rational fraction into partial fractions. We assume that the polynomials $P(x)$ and $Q(x)$ have *real* coefficients, the coefficient in x^n in the polynomial $Q(x)$ being equal to 1 (this can always be achieved by dividing the numerator and the denominator by the coefficient in x^n).

According to Sec. 75, I, the polynomial $Q(x)$ (the denominator of the given integrand rational fraction) can be written as a product of linear and quadratic factors with real coefficients:

$$Q(x) = (x - \alpha)^k \dots (x^2 + px + q)^t \dots \quad (*)$$

where α is a k -fold real root of the equation $Q(x) = 0$, and the quadratic equation $x^2 + px + q = 0$ has conjugate complex roots (that is $p^2 - 4q < 0$) which are t -fold conjugate complex roots of the same equation. When integrating the proper rational fraction $\frac{P(x)}{Q(x)}$ we suppose that the factorization $(*)$ of the polynomial $Q(x)$ into *real* factors is known or, which is the same, that all the roots of the equation $Q(x) = 0$ are known. Here we are not going to deal with the problem of finding the roots of the algebraic equation $Q(x) = 0$, and in what follows it will be assumed that factorization $(*)$ has been determined. It turns out that the fraction $\frac{P(x)}{Q(x)}$ can be represented as a decomposition into *partial fractions* of the form

$$\frac{A_l}{(x - \alpha)^l} \quad \text{and} \quad \frac{B_l x + C_l}{(x^2 + px + q)^l}$$

where A_l , B_l , C_l are constants, namely:

In the decomposition of the fraction $\frac{P(x)}{Q(x)}$ into partial fractions to every factor $(x - \alpha)^k$ of factorization $(*)$ of the denominator $Q(x)$ there corresponds a sum of k partial fractions of the form

$$\frac{A_k}{(x - \alpha)^k} + \frac{A_{k-1}}{(x - \alpha)^{k-1}} + \dots + \frac{A_1}{x - \alpha}$$

and to every factor $(x^2 + px + q)^t$ there corresponds a sum of t partial fractions of the form

$$\frac{B_t x + C_t}{(x^2 + px + q)^t} + \frac{B_{t-1} x + C_{t-1}}{(x^2 + px + q)^{t-1}} + \dots + \frac{B_1 x + C_1}{x^2 + px + q}$$

Here we shall not present the general proof of this assertion. The examples considered below demonstrate how the decomposition of a given fraction $\frac{P(x)}{Q(x)}$ into partial fractions can be obtained, that is, how all the unknown coefficients, A_t , B_t , C_t can be determined.

Thus, to the factorization (*) of the denominator $Q(x)$ there corresponds the decomposition

$$\begin{aligned} \frac{P(x)}{Q(x)} = & \frac{A_k}{(x-\alpha)^k} + \frac{A_{k-1}}{(x-\alpha)^{k-1}} + \dots + \frac{A_1}{x-\alpha} + \dots + \frac{B_t x + C_t}{(x^2 + px + q)^t} + \\ & + \frac{B_{t-1} x + C_{t-1}}{(x^2 + px + q)^{t-1}} + \dots + \frac{B_1 x + C_1}{x^2 + px + q} + \dots \end{aligned} \quad (**)$$

of the fraction $\frac{P(x)}{Q(x)}$ into partial fractions.

It follows that the integral of every rational fraction can be reduced to integrals of partial rational fractions of the indicated types, the latter being easily found as demonstrated by the examples below. These examples must be thoroughly studied as with their help we illustrate, without general considerations, the possible cases of integrating rational fractions.

Thus, *when applying the method of integration of rational fractions presented here we proceed from the factorization of the denominator of the given fraction $\frac{P(x)}{Q(x)}$, form decomposition (**) and replace the given integral by the sum of the integrals of the corresponding partial fractions.*

II. Examples.

(1) Take $\int \frac{x-3}{x^3-x} dx$. The decomposition of the integrand into partial fractions must be of the form

$$\frac{x-3}{x^2-x} = \frac{x-3}{x(x-1)(x+1)} = \frac{A}{x} + \frac{B}{x-1} + \frac{C}{x+1}$$

(the unknown coefficients can, of course, be denoted by any letters). First of all we must find the coefficients A , B and C .

Clearing this equations we obtain the relation

$$x-3 = A(x^2-1) + Bx(x+1) + Cx(x-1)$$

Since it is an identity, the coefficients in like powers of x on both sides must be equal (see Property 2 on page 247):

$$0 = A + B + C, \quad 1 = B - C, \quad -3 = -A$$

From this system of three equations with three unknowns we find

$$A=3, \quad B=-1, \quad C=-2$$

There is a still simpler method of determining the coefficients A, B , and C : putting in succession $x=0$, $x=1$, $x=-1$, we receive $-3=-A$, $-2=2B$ and $-4=2C$, that is, again

$$A=3, \quad B=-1, \quad C=-2$$

Thus, we obtain the identity

$$\frac{x-3}{x^3-x} = \frac{3}{x} - \frac{1}{x-1} - \frac{2}{x+1}$$

and therefore

$$\begin{aligned} \int \frac{x-3}{x^3-x} dx &= 3 \int \frac{dx}{x} - \int \frac{dx}{x-1} - 2 \int \frac{dx}{x+1} = \\ &= 3 \ln|x| - \ln|x-1| - 2 \ln|x+1| + C = \\ &= \ln \left| \frac{x^3}{(x-1)(x+1)^2} \right| + C \end{aligned}$$

(2) Let us find the integral $\int \frac{x-5}{x^3-3x^2+4} dx$. To this end we should factorize the denominator of the integrand fraction into factors. Observing that the denominator vanishes for $x=-1$ we divide it by $x+1$; in the quotient we obtain $x^2-4x+4=(x-2)^2$, and hence, the decomposition of the fraction into partial fractions must be of the form

$$\frac{x-5}{x^3-3x^2+4} = \frac{x-5}{(x+1)(x-2)^2} = \frac{A}{x+1} + \frac{B}{(x-2)^2} + \frac{C}{x-2}$$

Let us find the coefficients A, B, C . Reducing the fractions to a common denominator and clearing the equation of fractions we obtain

$$x-5 = A(x-2)^2 + B(x+1) + C(x+1)(x-2)$$

Here it is convenient to put $x=-1$, $x=2$ and $x=0$ which yields

$$-6=9A, \quad -3=3B \quad \text{and} \quad -5=4A+B-2C$$

whence

$$A=-\frac{2}{3}, \quad B=-1, \quad C=\frac{2}{3}$$

Consequently,

$$\begin{aligned} \int \frac{x-5}{x^3-3x^2+4} dx &= -\frac{2}{3} \int \frac{dx}{x+1} - \int \frac{dx}{(x-2)^2} + \frac{2}{3} \int \frac{dx}{x-2} = \\ &= -\frac{2}{3} \ln|x+1| + \frac{1}{x-2} + \frac{2}{3} \ln|x-2| + C \end{aligned}$$

(3) Consider $\int \frac{12}{x^4 + x^3 - x - 1} dx$. The denominator is readily factorized:

$$\begin{aligned} x^4 + x^3 - x - 1 &= x^3(x+1) - (x+1) = (x+1)(x^3-1) = \\ &= (x+1)(x-1)(x^2+x+1) \end{aligned}$$

Here the decomposition of the integrand into partial fractions has the form

$$\frac{12}{(x+1)(x-1)(x^2+x+1)} = \frac{A}{x+1} + \frac{B}{x-1} + \frac{Cx+D}{x^2+x+1}$$

The coefficients A, B, C, D are now found from the identity

$$\begin{aligned} 12 &= A(x-1)(x^2+x+1) + B(x+1)(x^2+x+1) + \\ &\quad + (Cx+D)(x+1)(x-1) \end{aligned}$$

On substituting four different numerical values of x , for instance, $x=1$, $x=-1$, $x=0$ and $x=2$, we obtain the system

$$\begin{aligned} 12 &= 6B, & 12 &= -A + B - D, \\ 12 &= -2A, & 12 &= 7A + 21B + 3(2C + D) \end{aligned}$$

which yields,

$$A = -6, \quad B = 2, \quad C = 4 \quad \text{and} \quad D = -4$$

Consequently,

$$\int \frac{12}{x^4 + x^3 - x - 1} dx = -6 \int \frac{dx}{x+1} + 2 \int \frac{dx}{x-1} + 4 \int \frac{x-1}{x^2+x+1} dx$$

The last integral on the right-hand side is found by splitting it into two integrals one of which contains in the numerator the derivative of the denominator:

$$\begin{aligned} \int \frac{x-1}{x^2+x+1} dx &= \frac{1}{2} \int \frac{2x+1-3}{x^2+x+1} dx = \\ &= \frac{1}{2} \int \frac{2x+1}{x^2+x+1} dx - \frac{3}{2} \int \frac{dx}{\left(x+\frac{1}{2}\right)^2 + \frac{3}{4}} = \\ &= \frac{1}{2} \ln(x^2+x+1) - \frac{3}{2} \frac{1}{\sqrt{\frac{3}{4}}} \arctan \frac{x+\frac{1}{2}}{\sqrt{\frac{3}{4}}} + C \end{aligned}$$

Finally,

$$\begin{aligned} \int \frac{12}{x^4 + x^3 - x - 1} dx &= -6 \ln|x+1| + 2 \ln|x-1| + \\ &\quad + 2 \ln(x^2+x+1) - 4\sqrt{3} \arctan \frac{2x+1}{\sqrt{3}} + C \end{aligned}$$

(4) Find the integral $\int \frac{x^2+x-1}{x(x^2+1)^2} dx$. The denominator of the integrand fraction has already been factorized; hence

$$\frac{x^2+x-1}{x(x^2+1)^2} = \frac{A}{x} + \frac{Bx+C}{(x^2+1)^2} + \frac{Dx+E}{x^2+1}$$

On clearing the fractions we obtain

$$x^2+x-1 = A(x^2+1)^2 + (Bx+C)x + (Dx+E)x(x^2+1)$$

Here it is convenient to equate the coefficients in like powers of x on the left-hand and right-hand sides. Beginning with the coefficients in x^4 we obtain the following system:

$$\begin{aligned} A+D &= 0, & E &= 0, & 2A+B+D &= 1 \\ C+E &= 1, & A &= -1 \end{aligned}$$

The solution of this system is

$$A = -1, \quad B = 2, \quad C = 1, \quad D = 1, \quad E = 0$$

Hence, we arrive at the identity

$$\frac{x^2+x-1}{x(x^2+1)^2} = \frac{-1}{x} + \frac{2x+1}{(x^2+1)^2} + \frac{x}{x^2+1}$$

and therefore,

$$\int \frac{x^2+x-1}{x(x^2+1)^2} dx = -\int \frac{dx}{x} + \int \frac{2x+1}{(x^2+1)^2} dx + \int \frac{x}{x^2+1} dx$$

The first and the third integrals on the right-hand side are known: they are equal, respectively, to $-\ln x$ and $\frac{1}{2} \ln(x^2+1)$.

Now we turn to the second integral:

$$\int \frac{2x+1}{(x^2+1)^2} dx = \int \frac{2x}{(x^2+1)^2} dx + \int \frac{dx}{(x^2+1)^2} = -\frac{1}{x^2+1} + \int \frac{dx}{(x^2+1)^2}$$

Let us denote the last integral on the right-hand side as I_2 :

$$I_2 = \int \frac{1}{(x^2+1)^2} dx$$

Here it is convenient to apply the following technique. We add x^2 to and simultaneously subtract it from the numerator of the integrand and break the integral into two integrals:

$$I_2 = \int \frac{x^2+1-x^2}{(x^2+1)^2} dx = \int \frac{dx}{x^2+1} - \int \frac{x^2}{(x^2+1)^2} dx$$

The second integral is integrated by parts:

$$\begin{aligned} \frac{x dx}{(x^2+1)^2} &= dv, \quad v = \frac{1}{2} \int \frac{2x dx}{(x^2+1)^2} = -\frac{1}{2} \frac{1}{x^2+1} \\ x &= u, \quad du = dx \end{aligned}$$

and

$$\int \frac{x^2 dx}{(x^2+1)^2} = -\frac{x}{2(x^2+1)} + \frac{1}{2} \int \frac{dx}{x^2+1}$$

On substituting this result into I_2 we find

$$I_2 = \frac{x}{2(x^2+1)} + \frac{1}{2} \int \frac{dx}{x^2+1} = \frac{x}{2(x^2+1)} + \frac{1}{2} \arctan x + C$$

Finally, the original integral is expressed in the form

$$\begin{aligned} \int \frac{x^2+x-1}{x(x^2+1)^2} dx &= -\ln x - \frac{1}{x^2+1} + \frac{x}{2(x^2+1)} + \frac{1}{2} \arctan x + \\ &+ \frac{1}{2} \ln(x^2+1) + C = \frac{x-2}{2(x^2+1)} + \ln \frac{\sqrt{x^2+1}}{x} + \frac{1}{2} \arctan x + C \end{aligned}$$

The correctness of the result can be easily checked: the derivative of the right-hand side is equal to the integrand.

The technique used here for finding I_2 can be applied to the general integral

$$I_k = \int \frac{dx}{(x^2+1)^k}$$

with an arbitrary integral exponent $k > 1$. If we write the identity

$$\frac{1}{(x^2+1)^k} = \frac{1}{(x^2+1)^{k-1}} - \frac{x^2}{(x^2+1)^k}$$

and transform the integral of the second summand on the right-hand side with the aid of integration by parts, as was done above, we obtain the equality

$$I_k = \frac{x}{2(k-1)(x^2+1)^{k-1}} + \frac{2k-3}{2(k-1)} I_{k-1}$$

This is the so-called *recurrence formula* for finding I_k . Replacing k by $k-1$ in this formula we obtain the expression of the integral I_{k-1} in terms of I_{k-2} , etc. Finally we arrive at the known integral

$$I_1 = \int \frac{dx}{x^2+1} = \arctan x + C$$

which completes the integration. As an instance, we leave it to the reader to verify that

$$I_4 = \frac{x}{6(x^2+1)^3} + \frac{5x}{24(x^2+1)^2} + \frac{5x}{16(x^2+1)} + \frac{5}{16} \arctan x + C$$

In conclusion it should be stressed once again that, despite some purely algebraic difficulties, the integral of any rational function can always be found; the above examples show that it is expressible in terms of rational functions, logarithms and arctangents, and thus is an *elementary* function. This remark is essentially important

because in many cases even integrals of very simple functions are not elementary functions (see Sec. 85, III), and any attempt to reduce them to tabular integrals is a mere waste of time.

Therefore, very often in integration processes emphasis is mainly laid upon finding a change of variable which reduces the given integral to an integral of a rational function; then the further course of integration becomes clear. We say that such a change *rationalizes* the integral. General examples of this kind are given in Secs. 83 and 84.

In what follows the symbol $R(\alpha, \beta, \gamma, \dots)$ will be used to denote a *rational* expression in the variables $\alpha, \beta, \gamma, \dots$, which means that on $\alpha, \beta, \gamma, \dots$ only arithmetical operations are performed. It is obvious that if each of the variables $\alpha, \beta, \gamma, \dots$ is a rational function of a variable u then the whole expression $R(\alpha, \beta, \gamma, \dots)$ is also a rational function of u .

83. Integrating Some Simple Irrational Functions.

I. Let us consider integrals of the form

$$\int R(x, \sqrt[k]{x}, \sqrt[m]{x}, \dots) dx \quad (*)$$

where the integrand is a *rational function* with respect to the variable of integration x and the radicals of x entering into it. We denote by n the *least common multiple* of all the exponents $k, m, \dots$; then all the quotients $\frac{n}{k} = r_1, \frac{n}{m} = r_2, \dots$ are integers.

The change of variables

$$x = u^n, \quad dx = nu^{n-1} du$$

reduces integral $(*)$ to an integral of a *rational function* (i.e. $(*)$ is *rationalized* by this change).

Indeed, here integral $(*)$ assumes the form

$$\int R(u^n, u^{r_1}, u^{r_2}, \dots) nu^{n-1} du$$

and, according to what was said above, the integrand is a rational function of u .

Example. Let us take $\int \frac{dx}{\sqrt{x+\sqrt[3]{x}}}$. We put $x = u^6$; then $\sqrt{x} = u^3$, $\sqrt[3]{x} = u^2$, $dx = 6u^5 du$ and

$$\int \frac{dx}{\sqrt{x+\sqrt[3]{x}}} = 6 \int \frac{u^5}{u^3+u^2} du = 6 \int \frac{u^3}{u+1} du$$

The rest is quite simple: on separating the entire part and integrating we obtain

$$6 \int \frac{u^3}{u+1} du = 6 \left(\frac{u^3}{3} - \frac{u^2}{2} + u - \ln|u+1| \right) + C$$

after which we return to the original variable:

$$\int \frac{dx}{\sqrt{x+\sqrt[3]{x}}} = 2\sqrt{x} - 3\sqrt[3]{x} + 6\sqrt[6]{x} - 6\ln|\sqrt[6]{x}+1| + C$$

A similar technique is used when all the radicands are equal to one and the same linear-fractional function $\frac{ax+b}{px+q}$, that is the integral has the form

$$\int R\left(x, \sqrt[k]{\frac{ax+b}{px+q}}, \sqrt[m]{\frac{ax+b}{px+q}}, \dots\right) dx$$

If this case the integral is rationalized by the substitution

$$\frac{ax+b}{px+q} = u^n, \quad x = \frac{qu^n - b}{a - pu^n}, \quad dx = \frac{aq - bp}{(a - pu^n)^2} nu^{n-1} du$$

If $p=0$, $q=1$, i.e. the radicand is a linear function, all the transformations are essentially simplified.

II. Let us consider some integrals involving an irrational expression $\sqrt{ax^2+bx+c}$.

(1) We begin with $\int \frac{dx}{\sqrt{ax^2+bx+c}}$. Taking the factor $\frac{1}{\sqrt{a}}$ (if $a > 0$) or $\frac{1}{\sqrt{-a}}$ (if $a < 0$) outside the symbol of integration we reduce the integral to the form

$$\int \frac{dx}{\sqrt{x^2+px+q}} \quad \text{or} \quad \int \frac{dx}{\sqrt{-x^2+px+q}}$$

Completing the squares we obtain

$$\int \frac{dx}{\sqrt{\left(x+\frac{p}{2}\right)^2 + \frac{4q-p^2}{4}}} \quad \text{or} \quad \int \frac{dx}{\sqrt{\frac{4q+p^2}{4} - \left(x-\frac{p}{2}\right)^2}}$$

According to the solution of Example 9 in Sec. 80, the first integral is expressible in terms of logarithms and the second integral reduces to the arcsine for $4q+p^2 > 0$ (if $4q+p^2 \leq 0$ in the second integral, the integrand assumes only imaginary values).

(2) Now let us consider the integral $\int \frac{dx}{(x-\alpha)\sqrt{ax^2+bx+c}}$. It can be easily verified that the substitution $x-\alpha = \frac{1}{u}$ reduces this integral to an integral of form (1).

(3) Let us consider $\int \frac{Ax+B}{\sqrt{ax^2+bx+c}} dx$. This integral can be split into two by separating the part of the numerator proportional to

the derivative of the radicand; then one of the integrals is directly found (as an integral of a power function) whereas the second is an integral of form (1).

(4) Finally, let us find $\int \sqrt{ax^2+bx+c} dx$. On completing the square under the radical sign we reduce the integral to one of the three already known integrals (Sec. 81, II, Examples 4 and 5) of the form

$$\int \sqrt{1-x^2} dx, \quad \int \sqrt{x^2+1} dx \quad \text{and} \quad \int \sqrt{x^2-1} dx$$

84. Integrating Trigonometric Functions. Let us be given an expression which involves rationally only trigonometric functions. Since every trigonometric function is expressed rationally in terms of $\sin x$ and $\cos x$ the given expression can be regarded as a rational function of $\sin x$ and $\cos x$, i.e. it has the form $R(\sin x, \cos x)$.

By the substitution $u = \tan \frac{x}{2}$ the integral $\int R(\sin x, \cos x) dx$ is rationalized (i.e. it is reduced to an integral of a rational function).

For, according to the well-known trigonometric formulas, we have

$$\sin x = 2 \sin \frac{x}{2} \cos \frac{x}{2} = 2 \tan \frac{x}{2} \cos^2 \frac{x}{2} = \frac{2 \tan \frac{x}{2}}{1 + \tan^2 \frac{x}{2}}$$

$$\cos x = \cos^2 \frac{x}{2} - \sin^2 \frac{x}{2} = \cos^2 \frac{x}{2} \left(1 - \tan^2 \frac{x}{2} \right) = \frac{1 - \tan^2 \frac{x}{2}}{1 + \tan^2 \frac{x}{2}}$$

i.e.

$$\sin x = \frac{2u}{1+u^2}, \quad \cos x = \frac{1-u^2}{1+u^2}$$

From the equality $x = 2 \arctan u$ we have $dx = \frac{2}{1+u^2} du$. Thus, finally,

$$\int R(\sin x, \cos x) dx = \int R\left(\frac{2u}{1+u^2}, \frac{1-u^2}{1+u^2}\right) \frac{2}{1+u^2} du$$

where the integrand is on the right-hand side rational in u .

Examples. (1) Consider $\int \frac{dx}{\sin x}$ and $\int \frac{dx}{\cos x}$. Putting $u = \tan \frac{x}{2}$ we obtain

$$\int \frac{dx}{\sin x} = \int \frac{1+u^2}{2u} \frac{2 du}{1+u^2} = \int \frac{du}{u} = \ln |u| + C = \ln \left| \tan \frac{x}{2} \right| + C$$

$$\int \frac{dx}{\cos x} = \int \frac{dx}{\sin \left(\frac{\pi}{2} + x \right)} = \ln \left| \tan \left(\frac{x}{2} + \frac{\pi}{4} \right) \right| + C$$

(2) Take $\int \frac{dx}{3+5\cos x}$. We put $u = \tan \frac{x}{2}$; then

$$\begin{aligned} \int \frac{dx}{3+5\cos x} &= \int \frac{2du}{(1+u^2) \left(3+5\frac{1-u^2}{1+u^2}\right)} = \\ &= \int \frac{du}{4-u^2} = \frac{1}{4} \ln \left| \frac{2+u}{2-u} \right| + C = \frac{1}{4} \ln \left| \frac{2+\tan \frac{x}{2}}{2-\tan \frac{x}{2}} \right| + C \end{aligned}$$

Generally, this method is very convenient for computing integrals of the form $\int \frac{dx}{a\cos x + b\sin x + c}$.

The integration of functions of the form $R(\sin x, \cos x)$ with the aid of the substitution $u = \tan \frac{x}{2}$ is sure to give the desired result, but it is because of its generality that this method may not be the best from the point of view of brevity and simplicity of the transformations involved.

If the expression of the function R only involves *even* powers of $\sin x$ and $\cos x$ the substitution $u = \tan x$ is more convenient. Then $\sin^2 x = \frac{u^2}{1+u^2}$, $\cos^2 x = \frac{1}{1+u^2}$, $dx = \frac{du}{1+u^2}$ and the integral is readily rationalized.

(3) We have $\int \frac{dx}{\sin^2 x - 3\cos^2 x} = \int \frac{du}{(1+u^2) \left(\frac{u^2}{1+u^2} - \frac{3}{1+u^2} \right)} =$
 $= \frac{1}{2\sqrt{3}} \ln \left| \frac{u - \sqrt{3}}{u + \sqrt{3}} \right| + C = \frac{1}{2\sqrt{3}} \ln \left| \frac{\tan x - \sqrt{3}}{\tan x + \sqrt{3}} \right| + C$ where $u = \tan x$.
 The same substitution is applied to integrals of the form $\int R(\tan x) dx$.

(4) For instance, $\int \frac{dx}{1+2\tan x} = \int \frac{du}{(1+u^2)(1+2u)}$. On finding the integral of the rational fraction and returning to the original variable we get

$$\int \frac{dx}{1+2\tan x} = \frac{1}{5} [x + 2 \ln |\cos x + 2 \sin x|] + C.$$

It is often advantageous to put $u = \sin x$ or $u = \cos x$. As an example, let us consider integrals of the form

$$\int \sin^m x \cos^n x dx$$

where m and n are integers. The techniques for finding them demonstrated below make it possible to avoid integrating rational fractions.

(a) If $m > 0$ and is *odd*, the substitution $\cos x = u$ immediately reduces the integral to a sum of integrals of power functions; if $n > 0$ and is *odd*, the substitution $\sin x = u$ yields the same result.

(5) For instance, $\int \frac{\sin^5 x}{\cos^2 x} dx = \int \frac{\sin^4 x}{\cos^2 x} \sin x dx = -\int \frac{(1-u^2)^2}{u^2} du$, where $u = \cos x$.

On breaking the integral into a sum of integrals, integrating and returning to the variable x , we obtain

$$\int \frac{\sin^5 x}{\cos^2 x} dx = \frac{1}{\cos x} + 2 \cos x - \frac{\cos^3 x}{3} + C$$

(b) If both exponents m and n are *positive* and *even*, the result is readily obtained by means of trigonometric transformations with multiple arguments.

(6) For instance, $\int \sin^2 x \cos^4 x dx = \int \left(\frac{1 - \cos 2x}{2} \right) \left(\frac{1 + \cos 2x}{2} \right)^2 dx =$
 $= \frac{1}{8} \int \sin^2 2x (1 + \cos 2x) dx = \frac{1}{8} \int \frac{1 - \cos 4x}{2} dx + \frac{1}{16} \int \sin^2 2x \cos 2x dx$
 $\times (2x) = \frac{1}{16} x - \frac{1}{64} \sin 4x + \frac{1}{48} \sin^3 2x + C$

(The substitution $\tan x = u$ would lead to more lengthy calculations.)

(c) If both exponents m and n are *negative* and their sum is *even*, the substitution $u = \tan x$ (or $u = \cot x$) again reduces the integral to a sum of integrals of power functions. Here is an example:

$$(7) \int \frac{dx}{\sin^3 x \cos x} = \int \frac{du}{(1+u^2) \frac{u^3}{(1+u^2)^{3/2}} \frac{1}{(1+u^2)^{1/2}}} = \int \frac{1+u^2}{u^3} du =$$

$$= -\frac{1}{2u^2} + \ln |u| + C = -\frac{1}{2 \tan^2 x} + \ln |\tan x| + C$$

The same result is obtained if we artificially represent unity in the numerator as $(\sin^2 x + \cos^2 x)^k$ where $2k = |m+n| - 2$.

(8) Consider one more example:

$$\int \frac{dx}{\sin^6 x} = \int \frac{(\sin^2 x + \cos^2 x)^2}{\sin^6 x} dx = \int \frac{dx}{\sin^2 x} + 2 \int \frac{\cot^2 x}{\sin^2 x} dx + \int \frac{\cot^4 x}{\sin^2 x} dx =$$

$$= -\cot x - \frac{2}{3} \cot^3 x - \frac{1}{5} \cot^5 x + C$$

(d) If one of the exponents is equal to zero while the other is *negative* and *odd*, the general substitution $u = \tan \frac{x}{2}$ again leads to integrating power functions. This is illustrated by the examples:

$$(9) \int \frac{dx}{\sin^3 x} = \int \frac{2 du}{(1+u^2) \left(\frac{2u}{1+u^2} \right)^3} = -\frac{1}{8u^2} + \frac{1}{2} \ln |u| + \frac{u^2}{8} + C =$$

$$= -\frac{1}{8 \tan^2 \frac{x}{2}} + \frac{1}{2} \ln \left| \tan \frac{x}{2} \right| + \frac{\tan^2 \frac{x}{2}}{8} + C$$

$$(10) \int \frac{dx}{\cos^3 x} = \int \frac{dx}{\sin^3 \left(x + \frac{\pi}{2} \right)}$$

It is readily observed that for other exponents m and n the integral is also reducible to a sum of integrals of the forms considered above; for instance, in the case $m = -2$, $n = -3$ we have

$$\begin{aligned} \int \frac{dx}{\sin^2 x \cos^3 x} &= \int \frac{\sin^2 x + \cos^2 x}{\sin^2 x \cos^3 x} dx = \int \frac{dx}{\cos^3 x} + \int \frac{\sin^2 x + \cos^2 x}{\sin^2 x \cos x} dx = \\ &= \int \frac{dx}{\cos^3 x} + \int \frac{dx}{\cos x} + \int \frac{\cos x}{\sin^2 x} dx \end{aligned}$$

The difficulties connected with the variety of methods are fully compensated by the simplicity of the final integration.

Integrals of the type $\int \sin^m x \cos^n x dx$ are also easily found with the help of *recurrence formulas* making it possible to reduce the large exponents m and n to smaller ones. In the table given at the end of the book these formulas are numbered as 64 and 65. We shall illustrate the method of deriving such formulas by an example. Let us introduce the notation

$$I_n = \int \cos^n x dx$$

We put

$$dv = \cos x dx, \quad u = \cos^{n-1} x$$

then

$$v = \sin x, \quad du = -(n-1) \cos^{n-2} x \sin x dx$$

and

$$\begin{aligned} I_n &= \sin x \cos^{n-1} x + (n-1) \int \cos^{n-2} x \sin^2 x dx = \\ &= \sin x \cos^{n-1} x + (n-1) \int \cos^{n-2} x dx - (n-1) \int \cos^n x dx \end{aligned}$$

whence

$$n I_n = \sin x \cos^{n-1} x + (n-1) I_{n-2}$$

and, consequently,

$$I_n = \frac{1}{n} \sin x \cos^{n-1} x + \frac{n-1}{n} I_{n-2}$$

This is the *recurrence formula* which we wanted to obtain. It makes it possible to complete the problem of finding the integral I_n for an arbitrary n .

It should be noted that the integration of a number of other functions, for instance, of the type $R(x, \sqrt{ax^2+bx+c})$, i.e. functions dependent rationally on x and $\sqrt{ax^2+bx+c}$, can be reduced to the integration of rational functions by means of a variety of transformations and substitutions. Here we shall not dwell on this question in more detail.

85. Conclusion. Using Tables of Integrals.

I. Practical integration. It is a comparatively rare case when the integration of a function can be carried out according to standard rules. But even when there exists a theoretical scheme for obtaining the final result of integration it does not necessarily provide the best and the most economical method. In most cases there are various ways of integration, and in each case there are some particular circumstances which suggest a technique that is most economical. The mastery of integration (as well as of any other mathematical operation) requires not only a thorough knowledge of *how* a given integral can be ultimately found but also the skill in finding it *with the minimum of time and effort*.

To illustrate what has been said we shall consider a simple example. Let us take the integral

$$\int \frac{x^3}{(x+1)^4} dx$$

Here it would be unreasonable to apply the general rule (the decomposition of the given fraction into partial fractions). It is readily seen that the substitution $u=x+1$ simplifies the problem immediately:

$$\begin{aligned} \int \frac{x^3}{(x+1)^4} dx &= \int \frac{(u-1)^3}{u^4} du = \int \frac{du}{u} - 3 \int \frac{du}{u^2} + 3 \int \frac{du}{u^3} - \int \frac{du}{u^4} = \\ &= \ln|x+1| + \frac{3}{x+1} - \frac{3}{2(x+1)^2} + \frac{1}{3(x+1)^3} + C \end{aligned}$$

The resourcefulness and skill are eventually gained in solving a sufficiently large number of examples.

II. Using tables of integrals. The integration methods we have studied consist in transformations reducing the given integral to an integral already known, i.e. to one that can be found in the *table of integrals*. Up to now we have used only the *basic* (the most concise) table of integrals. It is obvious that the more complete the table of integrals we use, the easier the integration.

For practical purposes we usually use various reference books and tables of most commonly occurring integrals. At the end of the book there is a table of integrals which is not very extensive but sufficient for finding integrals in most cases dealt with in practice. The two examples below demonstrate its application.

(1) Let us consider the integral $\int \frac{x^2-3x+1}{\sqrt{-1+4x-2x^2}} dx$. We do not find it in our table, but it can be easily reduced to tabular integrals. First of all, the quadratic trinomial under the radical $-1+4x-2x^2$ is brought to the form of the quadratic binomial:

$$-1+4x-2x^2 = -[1+2(x^2-2x)] = 1-2(x-1)^2$$

Now we make the change of the variable of integration $u = x-1$ in the integral, i.e. put $x = u+1$, which leads to

$$\begin{aligned} \int \frac{x^2-3x+1}{\sqrt{-1+4x-2x^2}} dx &= \int \frac{(u+1)^2-3(u+1)+1}{\sqrt{1-2u^2}} du = \\ &= \int \frac{u^2}{\sqrt{1-2u^2}} du - \int \frac{u}{\sqrt{1-2u^2}} du - \int \frac{du}{\sqrt{1-2u^2}} \end{aligned}$$

Each of the three integrals on the right-hand side of the last equality can be found in the table (see, respectively Formulas 33, 31 and 29). Taking advantage of these formulas for $a = \sqrt{2}$, $b = 1$ we obtain

$$\begin{aligned} \int \frac{x^2-3x+1}{\sqrt{-1+4x-2x^2}} dx &= \left[\frac{1}{4\sqrt{2}} (-\sqrt{2u}\sqrt{1-2u^2} + \arcsin \sqrt{2u}) \right] + \\ &+ \left[\frac{1}{2}\sqrt{1-2u^2} \right] - \left[\frac{1}{\sqrt{2}} \arcsin \sqrt{2u} \right] + C = \\ &= \frac{1}{4}(2-u)\sqrt{1+2u^2} - \frac{3}{4\sqrt{2}} \arcsin \sqrt{2u} + C \end{aligned}$$

On returning to the original variable x we arrive at the final result:

$$\begin{aligned} \int \frac{x^2-3x+1}{\sqrt{-1+4x-2x^2}} dx &= \frac{1}{4}(3-x)\sqrt{-1+4x-2x^2} - \\ &- \frac{3}{4\sqrt{2}} \arcsin \sqrt{2}(x-1) + C \end{aligned}$$

(2) Let us find $\int \sin^4 x \cos^2 x dx$. By Formula 64, for $m=4$, $n=2$ we get (see the second line of this formula):

$$\int \sin^4 x \cos^2 x dx = \frac{1}{6} \sin^5 x \cos x + \frac{1}{6} \int \sin^4 x dx$$

The integral on the right-hand side of this equality can again be found by Formula 64 for $m=4$, $n=0$ (see the first line of the formula):

$$\int \sin^4 x dx = -\frac{1}{4} \sin^3 x \cos x + \frac{3}{4} \int \sin^2 x dx$$

Furthermore, for $m=2$, $n=0$ we obtain

$$\int \sin^2 x \, dx = -\frac{1}{2} \sin x \cos x + \frac{1}{2} \int dx = -\frac{1}{2} \sin x \cos x + \frac{1}{2} x + C$$

Hence, finally,

$$\begin{aligned} \int \sin^4 x \cos^2 x \, dx = & \frac{1}{6} \sin^5 x \cos x - \frac{1}{24} \sin^3 x \cos x - \frac{1}{16} \sin x \cos x + \\ & + \frac{1}{16} x + C \end{aligned}$$

III. On expressing integrals in terms of elementary functions. In contrast to differentiation, the operation of integration of an elementary function by far not always leads to an elementary function. It can be proved rigorously that there exist elementary functions whose integrals are inexpressible in terms of *basic* elementary functions. Such functions are said to be *nonintegrable in elementary functions* (*nonintegrable in finite form*).

For instance, the integrals

$$\begin{aligned} \int \frac{dx}{\sqrt{1+x^2}}, \quad \int \frac{e^x}{x} dx, \quad \int \frac{\sin x}{x} dx, \quad \int \frac{\cos x}{x} dx, \\ \int \frac{dx}{\ln x} \quad \text{and} \quad \int e^{-x^2} dx \end{aligned}$$

cannot be represented by any elementary functions.

One should distinguish between the question of *existence* of a desired antiderivative and the possibility of expressing it with the aid of *elementary functions*. The integrals written above exist, but the class of all elementary functions which we use is insufficient for expressing the integrals as finite combinations of these functions.

Let us discuss this question in more detail. Suppose we limit ourselves to using all the basic elementary functions except logarithmic. Then many integrals whose expressions we know should be regarded as "inexpressible in elementary functions". For instance, the integrals $\int \frac{dx}{x}$, $\int \frac{dx}{x^2-1}$, cannot now be expressed "in finite form". For among the remaining basic elementary functions there is no function whose derivative is equal to $\frac{1}{x}$ or $\frac{1}{x^2-1}$ (since we cannot use the logarithmic function). But if the logarithmic function is included it becomes possible to find these integrals in finite form. But there is nothing surprising in the fact that some other integrals remain nonintegrable in elementary functions. To make them integrable it is necessary to *extend* the class of basic functions which we agree to use. This is exactly what is done in mathematical analysis: the *nonelementary* functions determined by the most important integrals inexpressible in terms of elementary functions are thoroughly investigated and tabulated.

These new functions extend the variety of our techniques and make it possible to express the integrals of a number of formerly non-integrable functions.

§ 2. Definite Integral

86. Some Problems of Geometry and Physics. There are various problems leading to the notion of the *definite integral*: determining the *area* of a plane figure, computing the *work* of a variable force, finding the *distance* travelled by a body with a given velocity, and many other problems. We shall discuss these problems in succession. At first glance it may seem that the methods of their solution are irrelevant to the foregoing sections devoted to indefinite integration. But later on, as the properties of the definite integral will be investigated, a close connection of these methods with the indefinite integral will be elucidated. This interrelation creates a basis for practical solution of these problems.

1. Area of a curvilinear trapezoid. The theory of area is based on two propositions:

(1) the area of a figure consisting of several figures is equal to the sum of the areas of the constituent figures;

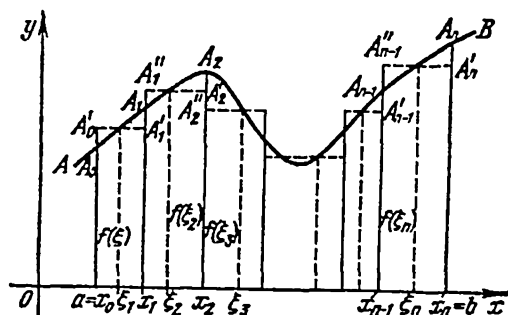


Fig. 103

(2) the area of a rectangle is equal to the product of its dimensions.

In elementary geometry, on the basis of these propositions, we find the area of a triangle and hence the area of a polygon since every polygon can be broken up into triangles. By considering the limit of the areas of inscribed and circumscribed regular polygons of the circle we can find its area as well. But here we make use of special geometrical properties of the figure (the circle), which may be inapplicable to other cases.

We shall begin with defining the area of a *curvilinear trapezoid*, i.e. a figure bounded by the x -axis, by the graph of a continuous function $y=f(x)$ and by two straight lines $x=a$ and $x=b$ (Fig. 103);

the interval $[a, b]$ of the x -axis will be called the *base* of the curvilinear trapezoid¹.

We shall assume that $f(x) > 0$ in the interval $[a, b]$, i.e. that the trapezoid lies over Ox .

Let us divide the base of the trapezoid (the interval $[a, b]$) into n subintervals

$$[x_0, x_1], [x_1, x_2], \dots, [x_{n-1}, x_n]$$

where $a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b$ and draw through the points of division of the interval $[a, b]$ the straight lines parallel to the axis of ordinates. Then the given curvilinear trapezoid is split up into n smaller trapezoids (Fig. 103).

Next we take in each subinterval an arbitrary point; denoting these points by $\xi_1, \xi_2, \dots, \xi_n$ we can write

$$x_0 \leq \xi_1 \leq x_1, \quad x_1 \leq \xi_2 \leq x_2, \quad \dots, \quad x_{n-1} \leq \xi_n \leq x_n$$

(Some of the chosen points ξ_i and even all of them may coincide with the points of division of the interval $[a, b]$ into the subintervals.)

Through the points $\xi_1, \xi_2, \dots, \xi_n$ let us draw the straight lines parallel to Oy to intersect the curve $y = f(x)$; the segments of these straight lines thus obtained are the ordinates of the corresponding points of the graph; they are, respectively, equal to $f(\xi_1), f(\xi_2), \dots, f(\xi_n)$.

On the subintervals as bases we now construct n rectangles with altitudes equal, respectively, to the values of the function $f(x)$ at the points ξ_i (Fig. 103). This results in the appearance of a "step-like" figure bounded above by the polygonal line $A_0A_1A_2A_3 \dots A_n$ with n segments.

It is natural to consider the area s_n of this step figure as an approximation to the area s of the given curvilinear trapezoid; the greater n and the smaller the lengths of the subintervals, the greater is the accuracy of the approximation. In other words, we proceed from the visual idea that as the number of the rectangles increases and they become narrower, the polygonal line $A_0A_1 \dots$ becomes closer to the given line $A_0A_1 \dots$ and the area s_n of the corresponding n -step figure gives a better approximation to the area of the curvilinear trapezoid. Hence, it is clear that the area s of the curvilinear trapezoid bounded above by the line $A_0A_1A_2 \dots A_n$ should be defined as the limit to which the variable area of the step figure bounded above by the polygonal line $A_0A_1A_2 \dots A_n$ tends when the number n increases indefinitely and the maximum length of the subintervals tends to zero.

¹ Note that according to elementary geometry it would be more natural to apply the term "base" to the parallel line segments $x=a$ and $x=b$ and consider the interval $[a, b]$ of the x -axis as the altitude of the trapezoid. This discrepancy in the terminology is of course inessential for our further aims.

It should be noted that the points $x_1, x_2, \dots, x_{n-1}$ can be chosen arbitrarily but it is required that the maximum length of the subintervals should tend to zero as the number n of the subintervals tends to infinity. For, if this is not stipulated, it may happen that the step figure does not approach the curvilinear trapezoid indefinitely. Indeed, let us increase the number of points of division of the base so that one of the subintervals, say (x_{n-1}, b) (Fig. 103), remains invariable. Then the polygonal line can indefinitely approach the arc A_0A_{n-1} of the given line $y=f(x)$ but it will not approach the arc $A_{n-1}A_n$, and, in this process, the invariable part of the trapezoid $x_{n-1}A_{n-1}A_nb$ will not be sufficiently accurately replaced by the rectangles. If we require that the maximum length of the subintervals should tend to zero this implies that the number n of subintervals increases indefinitely.

Now we can write the expression for s_n . Since the entire figure consists of n rectangles and the area of the rectangle corresponding to the subinterval $[x_{i-1}, x_i]$ is obviously equal to $f(\xi_i)(x_i - x_{i-1})$, we have

$$s_n = f(\xi_1)(x_1 - x_0) + f(\xi_2)(x_2 - x_1) + \dots \\ \dots + f(\xi_i)(x_i - x_{i-1}) + \dots + f(\xi_n)(x_n - x_{n-1})$$

All the members of this sum have the same form: they differ from each other only in the values of the subscript the independent variable is supplied with.

For our further aims it is convenient to introduce the *summation sign* Σ (the Greek capital letter "sigma") commonly used for contracted notation of sums. To this end we write

$$s_n = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \quad (*)$$

Generally, this means that the given expression (the general term) dependent on an index assuming integral values has to be summed with respect to that index as the index ranges from the value indicated below the letter Σ to the one indicated above it.

Now, according to the *definition* of the area s , we have

$$s = \lim s_n = \lim \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1})$$

As was said above, we assume that the limit is taken on condition that the length of the greatest subinterval, i.e. $\max(x_i - x_{i-1})$, tends to zero.

After the area of a curvilinear trapezoid has been defined, we are able to find the areas of many other figures. As a rule, a given plane figure can be broken up into a number of curvilinear trapezoids, and thus the sought-for area can be found as the

algebraic sum of the areas of the constituent curvilinear trapezoids. For instance, the area of the figure shown in Fig. 104 can be represented by the following algebraic sum:

$$\text{area } AA'C'C - \text{area } BB'C'C + \text{area } BB'D'D - \text{area } AA'D'D$$

II. The work of a variable force. We proceed now to an important physical problem of computing the work of a force. As will be seen, the solution of the problem of computing the work is completely analogous to that of determining the area of a trapezoid.

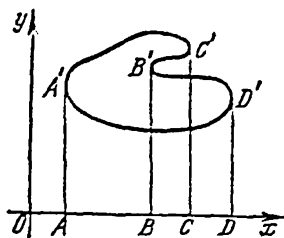


Fig. 104

Let a body move in a straight line under the action of a force, the direction of the force coinciding with the direction of motion. It is required to compute the work performed by the force as the body is moved from a position M to a position N (Fig. 105).

If the force remains constant along the whole path from M to N , the work is known to equal the product of the force by the path length. Let A denote the work, P the force and S the path length MN . Then¹

$$A = PS$$

Suppose, however, that the force varies along the path from M to N . At every point of the segment MN lying at a distance s

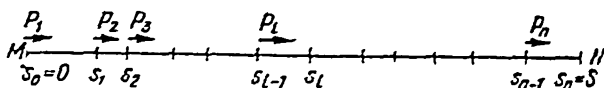


Fig. 105

from M the force acting upon the body assumes a certain value P corresponding to this point. This means that the force P is a function of the distance s , i.e. $P = f(s)$. How shall we define, in this case, the work performed when the body is moved from the point M to the point N ?

Let us divide the whole path MN , i.e. the interval of variation of the variable s , into n subintervals by points lying at distances

$$s_0 = 0, s_1, s_2, \dots, s_{l-1}, s_l, \dots, s_{n-1}, s_n = S$$

from the point M . Instead of the variable force P acting along the path MN we take another force P^* which preserves constant values in each of the subintervals, these values being equal to the values of the acting force P at arbitrarily fixed points of the

¹ If P is expressed in newtons and S in metres, A is expressed in joules.

subintervals. In the first subinterval $[s_0, s_1]$ this force is equal to $P_1 = f(\sigma_1)$ where $\sigma_1 \in [s_0, s_1]$; in the second subinterval $[s_1, s_2]$ it is equal to $P_2 = f(\sigma_2)$ where $\sigma_2 \in [s_1, s_2]$; in the subinterval $[s_{i-1}, s_i]$ it is equal to $P_i = f(\sigma_i)$ where $\sigma_i \in [s_{i-1}, s_i]$, etc.

The work done over the whole path being equal to the sum of the works corresponding to the parts into which the path is broken up, the work performed by the force P^* is

$$A_n = f(\sigma_1)(s_1 - s_0) + f(\sigma_2)(s_2 - s_1) + \dots + f(\sigma_i)(s_i - s_{i-1}) + \dots \\ \dots + f(\sigma_n)(s_n - s_{n-1}) = \sum_{i=1}^n f(\sigma_i)(s_i - s_{i-1})$$

It appears natural to take A_n as an approximation to the sought-for work; the greater the number n and the smaller the parts into which the whole path MN is divided, the more accurate is the approximation. The work A is assumed *by definition* to be equal to the limit of A_n as $n \rightarrow \infty$ (as before, it is supposed that the length of the greatest subinterval tends to zero):

$$A = \lim A_n = \lim \sum_{i=1}^n f(\sigma_i)(s_i - s_{i-1})$$

III. Computing path length and mass. Let us briefly discuss two more physical problems: computing *path length* and determining *mass*. We shall solve these problems in exactly the same way as the previous ones.

(1) Let a material point be in a rectilinear motion and let $v = f(t)$ be its velocity. If $f(t) > 0$, the path s travelled by the point during the interval $[T_1, T_2]$ is defined as

$$s = \lim s_n = \lim \sum_{i=1}^n f(\tau_i)(t_i - t_{i-1})$$

where $t_0, t_1, t_2, \dots, t_{n-1}, t_n$ are points of division of the interval $[T_1, T_2]$ into n subintervals:

$$T_1 = t_0 < t_1 < t_2 < \dots < t_{n-1} < t_n = T_2$$

and

$$t_{i-1} \leq \tau_i \leq t_i \quad (i = 1, 2, \dots, n)$$

We again assume that $\max(t_i - t_{i-1}) \rightarrow 0$. Each term $f(\tau_i)(t_i - t_{i-1})$ of the sum expresses the path which the point would travel during the time interval $[t_{i-1}, t_i]$ if it moved with *constant* velocity equal to $f(\tau_i)$.

(2) Take a segment of a material line and suppose that we know its mass density δ at every point s , i.e. $\delta = f(s)$, the length s varying within the interval $[0, S]$.

Denoting by m the mass of the portion of the material line corresponding to $0 \leq s \leq S$ we put, by definition,

$$m = \lim m_n = \lim \sum_{i=1}^n f(\sigma_i) (s_i - s_{i-1})$$

where $s_0, s_1, s_2, \dots, s_{n-1}, s_n$ are points of division of the interval $[0, S]$ into n subintervals,

$$s_0 = 0 < s_1 < s_2 < \dots < s_{n-1} < s_n = S$$

and

$$s_{i-1} \leq \sigma_i \leq s_i \quad (i = 1, 2, \dots, n) \quad \text{and} \quad \max (s_i - s_{i-1}) \rightarrow 0$$

Each term $f(\sigma_i) (s_i - s_{i-1})$ of the sum expresses the mass which the portion $[s_{i-1}, s_i]$ of the line would carry if the density were constant along this portion and would be equal to $f(\sigma_i)$ (i.e. if the mass were distributed uniformly).

87. Definite Integral. Existence Theorem. We have seen in the previous section that the solutions of many important problems of geometry and physics (computing the area, work, path length and mass) lead to one and the same sequence of operations on given functions and their arguments. Since this sequence of operations is applied in various cases and is of great importance we should describe it formally (mathematically) irrespective of the concrete conditions of the problem. Then its application in every suitable case is taken for granted and does not require special justification.

Thus, let there be a bounded function $f(x)$ in a closed interval $[a, b]$; we do not require that the function should be positive, and so far we do not even assume that $f(x)$ is continuous.

The sequence of operations we speak of consists in the following:

(1) the interval $[a, b]$ in which the function $f(x)$ is defined is divided into n subintervals with the aid of the points

$$x_0 = a, x_1, x_2, \dots, x_{n-1}, x_n = b$$

where

$$a < x_1 < x_2 < \dots < x_{n-1} < b$$

(2) the value of the function $f(\xi_i)$ at a point $\xi_i \in [x_{i-1}, x_i]$ is multiplied by the length of the subinterval $x_i - x_{i-1}$, i.e. the product $f(\xi_i) (x_i - x_{i-1})$ is formed;

(3) the sum

$$\begin{aligned} I_n &= f(\xi_1) (x_1 - x_0) + f(\xi_2) (x_2 - x_1) + \dots + f(\xi_n) (x_n - x_{n-1}) = \\ &= \sum_{i=1}^n f(\xi_i) (x_i - x_{i-1}) \end{aligned}$$

of all these products is formed; if we designate $x_i - x_{i-1}$ by Δx_i then

$$I_n = \sum_{i=1}^n f(\xi_i) \Delta x_i \quad (\text{A})$$

The sum (A) is called an *integral sum*.

(4) The limit I of the integral sum I_n , as the length of the greatest subinterval tends to zero, is found; this means that $\max \Delta x_i \rightarrow 0$, and, consequently, $n \rightarrow \infty$. Thus,

$$I = \lim I_n = \lim \sum_{i=1}^n f(\xi_i) \Delta x_i$$

This limit (if it exists) is called the *definite integral* or simply the *integral* of the function $f(x)$ with respect to x over the interval $[a, b]$ and is denoted as

$$I = \int_a^b f(x) dx$$

(read "the integral of $f(x) dx$ from a to b "). Hence, by definition,

$$\int_a^b f(x) dx = \lim \sum_{i=1}^n f(\xi_i) \Delta x_i$$

An integral sum for a chosen number n of subintervals can be formed in different ways. For we not only break up the interval $[a, b]$ in an arbitrary manner but also choose an arbitrary point ξ_i in each subinterval. That is why the integral sum (A) is not an ordinary function of n : to one and the same value of n there may correspond an infinite number of various integral sums. Therefore the limit of the integral sum we spoke of is not an ordinary limit of a function or of a sequence, and we must formulate what should be understood as that limit:

The number I is called the limit of the integral sum (A), if, given an arbitrary $\varepsilon > 0$, there is a number $\delta > 0$ such that for any division of the interval $[a, b]$ into subintervals whose lengths are less than δ , i.e. for $\max \Delta x_i < \delta$, and for any choice of the intermediate points ξ_i there holds the inequality

$$\left| \sum_{i=1}^n f(\xi_i) \Delta x_i - I \right| < \varepsilon$$

It is this number $I = \int_a^b f(x) dx$ that is the definite integral of the function $f(x)$ over the interval $[a, b]$.

In what follows, when speaking of the limit of an integral sum, we shall always understand it in the sense stated here without writing any additional explanatory signs under the symbol of the limit. Briefly stated, the definition reads as follows:

Definition. The *definite integral* is the limit of the integral sum when the length of the greatest subinterval tends to zero (under the assumption that this limit exists).

The functions for which the definite integral exists are called *integrable*; the theorem stated at the end of this section shows that any continuous function is integrable.

The sign of integration $\int$ originates from an elongated letter *s*, the first letter of the Latin word *summa* sum. The expression $f(x)dx$ under the integral sign indicates the form of the terms entering into the integral sums with the summation index omitted (the absence of the index underlines that in the limiting process completing the computation of the integral the variable x runs over all the values lying in the interval $[a, b]$).

As in the case of the indefinite integral, the function $f(x)$ is called the *integrand*, the expression $f(x)dx$ is the *element of integration*, the interval $[a, b]$ is the *interval of integration* and the variable x is the *variable of integration*; the numbers a and b are called, respectively, the *lower* and the *upper limits*¹ of integration.

The process of forming the definite integral shows that the symbol $\int_a^b f(x)dx$ is a certain *number*. Its value only depends on the properties of the integrand and on the numbers a and b determining the interval of integration. The variable of integration x enters into the notation of the definite integral and nothing happens if it is denoted by any other letter, for instance, t or u :

$$\int_a^b f(x)dx = \int_a^b f(t)dt = \int_a^b f(u)du$$

The similarity in the notation of the definite and the indefinite integrals (the notation of the latter lacks only the limits of integration) emphasizes their close interrelation, though the definite integral is a *number* whereas the indefinite integral is the family of all *functions* which are antiderivatives of the given function. This interrelation will be established later (in Sec. 92), and at present we shall only note that it enables us to evaluate definite integrals with the aid of indefinite integration without resorting

¹ Here the word "limit" has nothing in common with the notion of the limit of a function.

to the technique of forming integral sums with further passage to the limit. This is the reason why we do not give examples of evaluating definite integrals with the aid of summation; this would involve considerable difficulties even in the simplest cases.

Notice should only be taken of the following obvious result:

$$\int_a^b dx = b - a \quad (*)$$

This is implied by the fact that any integral sum for the function $f(x) \equiv 1$ is equal to $b - a$:

$$\sum_{i=1}^n \Delta x_i = x_1 - x_0 + x_2 - x_1 + \dots + x_n - x_{n-1} = x_n - x_0 = b - a$$

Applying the definition of the definite integral to the four problems considered in Sec. 86 we can express the conclusions drawn there as follows:

(1) *the area of a curvilinear trapezoid is equal to the integral, of the ordinate of the curve bounding the trapezoid, taken over the base of the trapezoid:*

$$s = \int_a^b f(x) dx$$

(2) *the work performed by a force is equal to the integral of the force taken along the path covered:*

$$A = \int_0^s f(s) ds$$

(3) *the distance travelled by a material point is equal to the integral of the velocity with respect to time:*

$$s = \int_{\tau_1}^{\tau_2} f(t) dt$$

(4) *the mass distributed over a line is equal to the integral of the density along the line taken with respect to the arc length:*

$$m = \int_0^s f(s) ds$$

In the first three cases it is additionally stipulated that the integrand is *positive* (when computing the mass no stipulation is necessary because the density cannot be negative). Below (Sec. 89) we shall see how these conclusions should be modified when no such stipulations are made.

When speaking of the definite integral we did not discuss the question what functions it exists for, i.e. for what functions there exist the limits of their integral sums. The answer is given by the following theorem which is given without proof:

Theorem (On Existence of the Definite Integral). If the function $f(x)$ is continuous in the closed interval $[a, b]$, its integral sum tends to a definite limit as the length of the greatest subinterval tends to zero.

It should be stressed once again that this limit, i.e. the definite integral $\int_a^b f(x) dx$, does not depend on the method of breaking up the interval of integration into subintervals and on the choice of the intermediate points in them.

Integral sums formed for different divisions of the interval of integration and for different choices of the points ξ_i may substantially differ from each other. The remarkable theorem formulated above shows that for continuous functions the difference between these sums disappears as the number of points of division increases and the length of the greatest subinterval decreases and vanishes in the limit.

In what follows we assume, unless the contrary is stated, that the functions under consideration are continuous.

The existence theorem shows that to every function $f(x)$ continuous in a closed interval $[a, b]$ there corresponds a certain number equal to the definite integral $\int_a^b f(x) dx$. Here we come across a new type of functional dependence in which the continuous function $f(x)$ serves as an independent variable and the definite integral of $f(x)$ as a function.

Generally, if to every function from some class of functions there corresponds a certain number we say that there is a *functional* defined on that class of functions.

For instance, the definite integral $\int_a^b f(x) dx$ is a functional on the class of functions which are continuous in the closed interval $[a, b]$.

88. Simplest Properties of the Definite Integral.

Theorem I (On Integrating a Sum). The integral of a sum of a finite number of functions is equal to the sum of the integrals of these functions:

$$\int_a^b (u + v + \dots + w) dx = \int_a^b u dx + \int_a^b v dx + \dots + \int_a^b w dx$$

where $u, v, \dots, w$ are functions of the independent variable x .

Proof. Let us denote the integral on the left-hand side by I . According to the definition of the integral, we have

$$\begin{aligned} I &= \lim \sum_{i=1}^n (u_i + v_i + \dots + w_i) \Delta x_i = \\ &= \lim \left(\sum_{i=1}^n u_i \Delta x_i + \sum_{i=1}^n v_i \Delta x_i + \dots + \sum_{i=1}^n w_i \Delta x_i \right) \end{aligned}$$

where $u_i, v_i, \dots, w_i$ are, respectively, the values of the functions $u, v, \dots, w$ at some point of the interval $[x_{i-1}, x_i]$ and $\Delta x_i = x_i - x_{i-1}$. Taking advantage of the theorem on the limit of a sum¹ we write

$$I = \lim \sum_{i=1}^n u_i \Delta x_i + \lim \sum_{i=1}^n v_i \Delta x_i + \dots + \lim \sum_{i=1}^n w_i \Delta x_i$$

i.e.

$$I = \int_a^b u \, dx + \int_a^b v \, dx + \dots + \int_a^b w \, dx$$

which is what we set out to prove.

Theorem II (On Taking a Constant Factor Outside the Sign of Definite Integration). A constant factor in the integrand can be taken outside the integral sign:

$$\int_a^b c u \, dx = c \int_a^b u \, dx$$

where u is a function of the argument x and c is a constant.

Proof. Let us denote the integral on the left-hand side of the equality by I . According to the definition of the integral, we have

$$I = \lim \sum_{i=1}^n c u_i \Delta x_i$$

Taking the constant c outside the sign of the sum and then outside the sign of the limit we obtain

$$I = \lim c \sum_{i=1}^n u_i \Delta x_i = c \lim \sum_{i=1}^n u_i \Delta x_i = c \int_a^b u \, dx$$

which is what we set out to prove.

¹ The general theorems on limits of functions (including the theorem on the limit of a sum of functions) can be readily extended to limits of integral sums.

Theorems I and II imply that

$$\int_a^b (C_1 u_1 + \dots + C_n u_n) dx = C_1 \int_a^b u_1 dx + \dots + C_n \int_a^b u_n dx$$

where $C_1, \dots, C_n$ are constants and $u_1, \dots, u_n$ are functions of x .

89. Interchanging Limits in the Definite Integral and Splitting the Interval of Integration. Geometrical Meaning of the Integral. Up to now we have assumed that the lower limit of the integral is less than the upper or, as we say, the interval of integration

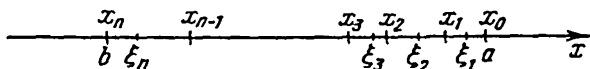


Fig. 106

is directed to the right. However, the general definition of the integral can be extended to embrace the case $a > b$ as well (for $a > b$ we say that the interval of integration is directed to the left). To this end we divide the interval of integration into subintervals by points $x_1, x_2, \dots, x_{n-1}$ (Fig. 106) where

$$a > x_1 > x_2 > \dots > x_{n-1} > b$$

and consider the *negative* differences $x_i - x_{i-1}$. Accordingly, every difference $x_i - x_{i-1}$, entering into the construction of the corresponding integral sum is no longer the length of the subinterval but differs from it in the sign.

If we now consider the two integrals

$$\int_a^b f(x) dx \quad \text{and} \quad \int_b^a f(x) dx$$

and form their integral sums for one and the same division of the interval into subintervals and for one and the same choice of the intermediate points, they obviously differ from each other *only in the sign*. Thus, we have arrived at the following theorem¹:

Theorem III (On Interchanging Limits of Integration). If the upper and the lower limits of the definite integral are interchanged the integral only changes its sign:

$$\int_a^b f(x) dx = - \int_b^a f(x) dx$$

¹ In Secs 88-92 the theorems are numbered in the consecutive order.

Bearing in mind Theorem III, in what follows we shall confine ourselves to the integral

$$\int_a^b f(x) dx$$

with $a < b$, for, if $a > b$, the change of the sign of the integrand reduces the given integral to an integral whose lower limit is less than the upper one.

It is also quite clear that if the upper and the lower limits of integration coincide, i.e. if $b = a$, the integral should be understood as equal to zero:

$$\int_a^a f(x) dx = 0$$

From the geometrical point of view this means that if the end points of the base of the trapezoid are made coincident, the trapezoid degenerates into a line segment of length $f(a)$ whose area must be considered equal to zero.

Theorem IV (On Splitting the Interval of Integration). If the interval of integration $[a, b]$ is split into two parts $[a, c]$ and $[c, b]$ then

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx \quad (*)$$

Proof. Since the limit of the integral sum does not depend on the way the interval $[a, b]$ is divided into parts (see the existence theorem in Sec. 87), we can divide the interval so that the point c is always a point of division. Then the integral sum can be represented in the form

$$\sum f(\xi_i) \Delta x_i = \sum_1 f(\xi_i) \Delta x_i + \sum_2 f(\xi_i) \Delta x_i$$

where the first sum on the right-hand side includes all the terms corresponding to the points of division of the interval $[a, c]$ and the second sum includes all the terms corresponding to the points of division of the interval $[c, b]$. The first and the second sums are, respectively, integral sums for the function $f(x)$ corresponding to the intervals $[a, c]$ and $[c, b]$. If the number of the points of division increases indefinitely and the length of the greatest subinterval tends to zero for the whole interval $[a, b]$ the same obviously takes place for the intervals $[a, c]$ and $[c, b]$; then the first sum tends to the integral from a to c and the second to the integral from c to b , and thus the required equality is obtained.

Equality (*) also holds in the case when the point c lies outside the interval $[a, b]$ to the right of it, i.e. $a < b < c$, or to the left

of it, i.e. $c < a < b$, on condition of course that the function $f(x)$ is continuous in $[a, c]$ or in $[c, b]$.

Indeed, let $a < b < c$. As has been proved, since b lies between a and c we have

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx$$

whence

$$\int_a^b f(x) dx = \int_a^c f(x) dx - \int_b^c f(x) dx$$

Now, interchanging the limits of integration in the second integral on the right-hand side we obtain

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

which is what we wanted to prove.

It can similarly be proved that equality (*) is valid for $c < a < b$.

From the theorem proved it follows that if $c_1, c_2, \dots, c_k$ are arbitrary points belonging to an interval of continuity $[a, b]$ of the function $f(x)$, then

$$\int_a^b f(x) dx = \int_a^{c_1} f(x) dx + \int_{c_1}^{c_2} f(x) dx + \dots + \int_{c_k}^b f(x) dx \quad (**)$$

Theorem IV expresses the so-called *additivity* property of the definite integral.

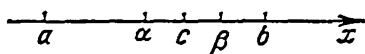


Fig. 107

Theorem V (On the Sign of the Integral). If the integrand retains sign in the interval of integration the integral is a number of the same sign as the function.

Proof. Let $f(x) \geq 0$ in the interval $[a, b]$ ($a < b$). Then in the integral sum $I_n = \sum_{i=1}^n f(\xi_i) \Delta x_i$ all the summands are nonnegative, and hence, $I_n \geq 0$; the limit of a nonnegative quantity cannot be negative, i.e. $\int_a^b f(x) dx \geq 0$.

It now remains to show that this integral cannot be equal to zero unless the function $f(x)$ is identically equal to zero. Let $f(c) > 0$ at a point c of the interval. Then, by the continuity of the function $f(x)$, there exists an interval $[\alpha, \beta]$ (Fig. 107) in

which the function remains positive (see Sec. 32). Let us denote by m its least value in the interval $[\alpha, \beta]$; then $f(x) \geq m > 0$ for $x \in [\alpha, \beta]$. Any integral sum for the interval $[\alpha, \beta]$ therefore satisfies the inequality

$$\sum_{k=1}^r f(\xi_k) \Delta x_k \geq \sum_{k=1}^r m \Delta x_k = m \sum_{k=1}^r \Delta x_k = m(\beta - \alpha)$$

But then for its limit we also have $\int_{\alpha}^{\beta} f(x) dx \geq m(\beta - \alpha) > 0$. By Theorem IV,

$$\int_a^b f(x) dx = \int_a^{\alpha} f(x) dx + \int_{\alpha}^{\beta} f(x) dx + \int_{\beta}^b f(x) dx$$

As has been proved, the first and the third integrals on the right-hand side of the equality are *nonnegative*, and the second, as we have just shown, is *positive*. Hence,

$$\int_a^b f(x) dx > 0$$

which is what we wanted to prove.

If the integrand in the interval of integration changes its sign the integral can be positive or negative or equal to zero.

When solving problems we should bear in mind the property of the definite integral implied by Theorem V. For instance, when computing the area of a curvilinear trapezoid with the aid of the integral it is necessary to take into consideration its disposition relative to the base. In the case when the trapezoid lies entirely above the x -axis the integral of the ordinate expresses the area; in the case when the trapezoid lies entirely below the x -axis the integral is negative and expresses the area of the trapezoid taken with the negative sign. Finally, when some parts of the trapezoid lie above the x -axis and some below it, it is necessary, when computing the area, to evaluate separately the integrals expressing the areas of the parts lying above the x -axis and the integrals expressing the areas of the parts lying below the x -axis and take the sum of their absolute values.

Geometrical meaning of the integral. Let us agree to ascribe the plus sign to the area of a curvilinear trapezoid situated above Ox and the minus sign to the area situated below Ox . Then, by Theorems IV and V, the definite integral of the function $f(x)$ is the *algebraic sum of the areas* of the curvilinear trapezoids, situated over and under Ox of which the given trapezoid is made up, each area being taken with the corresponding

sign. This sum is called the *algebraic area* of the whole curvilinear trapezoid.

Bearing this in mind, in what follows we can consider any definite integral

$$\int_a^b f(x) dx$$

irrespective of the concrete meaning of the variable of integration x and of the function $f(x)$, as the algebraic area of the curvilinear trapezoid with base $[a, b]$ bounded by the curve $y = f(x)$.

Let us return to the formula $s = \int_0^T f(t) dt$ expressing the distance travelled by a body in its rectilinear motion with velocity $v = f(t) > 0$ (see Sec. 86, III). If it is not assumed that $f(t) > 0$, the integral expresses the *algebraic sum of displacements* of the point, every displacement being taken with the sign plus or minus depending on the direction (positive or negative) in which it is reckoned. The difference between the distance and the algebraic sum of displacements is obvious; for instance, if the point is in a harmonic vibration, that is $s = A \sin \omega t$, then the path covered by it during the time interval $[0, T]$ where $T = \frac{2\pi}{\omega}$ is a period, is equal to $4|A|$, and the algebraic sum of displacements is equal to zero.

The formula expressing the work can be interpreted in a similar way.

90. Estimating Definite Integral. Mean Value Theorem. The Arithmetic Mean of a Function.

I. Estimation of definite integral. Let us determine upper and lower bounds for the value of a definite integral.

Theorem VI (On Estimating a Definite Integral). The value of a definite integral lies between the products of the least and the greatest values of the integrand by the length of the interval of integration, i.e.

$$m(b-a) \leq \int_a^b f(x) dx \leq M(b-a), \quad (a < b)$$

where m and M are, respectively, the least and the greatest values of the function $f(x)$ in the interval $[a, b]$:

$$m \leq f(x) \leq M$$

Proof. Let us take two functions $M - f(x)$ and $m - f(x)$, the former being nonnegative in the interval $[a, b]$, the latter nonpo-

sitive. By Theorem V (Sec. 89), we have

$$\int_a^b [M - f(x)] dx \geq 0 \quad \text{and} \quad \int_a^b [m - f(x)] dx \leq 0$$

Now, applying the theorems of Sec. 88 and formula (*) of Sec. 87 we obtain

$$M(b-a) \geq \int_a^b f(x) dx \quad \text{and} \quad m(b-a) \leq \int_a^b f(x) dx$$

which is what we set out to prove. From the proof of Theorem V it follows that unless the function $f(x)$ is constant the nonstrict inequalities can be replaced by the strict ones:

$$m(b-a) < \int_a^b f(x) dx < M(b-a)$$

The geometrical meaning of the inequalities we have just proved is the following: the area of a curvilinear trapezoid is greater than the area of the rectangle with altitude equal to the least ordinate of the trapezoid constructed on its base and is less than the area of the rectangle with altitude equal to the greatest ordinate of the trapezoid constructed on the same base (Fig. 108).

When finding bounds for an integral we *estimate* it. It may sometimes be difficult or even impossible to find the exact value of an integral, but by estimating it we learn, even if crudely, its approximate value. In mathematics such estimations are rather often used.

The bounds for an integral indicated in Theorem VI become more accurate as the interval of integration is made shorter and as the curve $y=f(x)$ approaches a straight line parallel to the x -axis.

Examples. (1) Let us estimate the integral

$$\int_0^2 \frac{5-x}{9-x^2} dx$$

With the aid of the well-known methods of differential calculus (Chapter IV) we find that the greatest and the least values of the integrand in the interval $[0, 2]$ are, respectively, equal to $y_{x=2} = 0.6$ and $y_{x=1} = 0.5$. Hence,

$$0.5(2-0) < \int_0^2 \frac{5-x}{9-x^2} dx < 0.6(2-0)$$

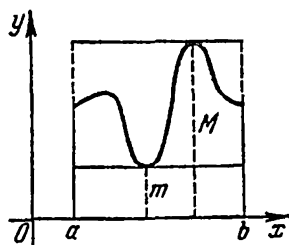


Fig. 108

i.e. the value of the integral lies between 1 and 1.2. If we assume that it is approximately equal to 1.1, the limiting absolute error is 0.1, and the relative error is $\frac{0.1}{1.1} 100 \approx 9\%$. Later we shall be able to find the exact value of the given integral which is equal to $\frac{4}{3} \ln 5 - \ln 3 \approx 1.047$.

(2) Let us estimate the integral

$$\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin x}{x} dx$$

It can be easily verified that the integrand decreases in the interval $\left[\frac{\pi}{4}, \frac{\pi}{2}\right]$ and, consequently,

$$\frac{\sin \frac{\pi}{2}}{\frac{\pi}{2}} \cdot \frac{\pi}{4} < \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin x}{x} dx < \frac{\sin \frac{\pi}{4}}{\frac{\pi}{4}} \cdot \frac{\pi}{4} \quad \text{i.e.} \quad \frac{1}{2} < \int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \frac{\sin x}{x} dx < \frac{1}{2} \sqrt{2}$$

Thus, the value of this integral lies between 0.5 and 0.71, which enables us to take the value 0.6 as its approximation accurate to within 0.1. More precise methods show that the integral is approximately equal to 0.62.

II. Generalized theorem on estimating definite integral. Integrating inequalities. There holds the following theorem more general than Theorem VI:

Theorem VII. If at every point x of an interval $[a, b]$ the inequalities

$$\psi(x) \leq f(x) \leq \varphi(x)$$

are fulfilled then

$$\int_a^b \psi(x) dx \leq \int_a^b f(x) dx \leq \int_a^b \varphi(x) dx, \quad a < b$$

This means that an inequality between functions implies an inequality of the same sense between their definite integrals, or, briefly speaking, that it is allowable to *integrate inequalities termwise*. On the other hand, simple geometrical considerations indicate that the differentiation of an inequality may lead to a senseless result. For instance, the inequality $f(x) < C = \text{const}$ does not, of course, imply that $f'(x) < 0$.

The proof of the theorem is immediately carried out by the application of Theorem V on the sign of an integral to the in-

equalities $f(x) - \varphi(x) \leq 0$ and $f(x) - \psi(x) \geq 0$. The equality signs in Theorem VII only appear when the functions coincide identically.

In particular, when $\varphi(x)$ is identically equal to M and $\psi(x)$ is identically equal to m we get Theorem VI.

With the aid of Theorem VII we can readily obtain an important inequality which will be of use in what follows. For any x we have

$$-|f(x)| \leq f(x) \leq |f(x)|$$

(If $f(x) > 0$ the right inequality turns into an equality and the left inequality is obvious; if $f(x) < 0$ these inequalities are proved similarly.) It follows that

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx$$

that is,

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

Thus, the absolute value of the definite integral does not exceed the integral of the absolute value of the integrand function. Let the reader elucidate the geometrical meaning of this inequality. We remind the reader that a similar proposition applies to sums: the absolute value of a sum does not exceed the sum of the absolute values of the summands.

III. Mean value theorem. The definite integral possesses the following important property:

Theorem VIII (The Mean Value Theorem). Let $f(x)$ be a continuous function in a closed interval $[a, b]$. Then there exists at least one point $x = \xi$ inside this interval such that

$$\frac{\int_a^b f(x) dx}{b-a} = f(\xi) \quad (*)$$

Proof. If the function $f(x)$ is constant formula (*) is obvious, ξ being any point of the interval $[a, b]$. Suppose $f(x)$ is nonconstant; then, by Theorem VI, we have

$$m < \frac{\int_a^b f(x) dx}{b-a} < M$$

and, hence,

$$\frac{\int_a^b f(x) dx}{b-a} = \mu$$

where μ is a number lying between the least value m and the greatest value M of the function $f(x)$ in the interval $[a, b]$.

The well-known property of continuous functions (see Sec. 35) implies that there are at least two points of the interval $[a, b]$ at which, respectively, the function $f(x)$ assumes the values m and M and also a point lying between them where it assumes the intermediate value μ . Hence, there exists a point $\xi \in (a, b)$ at which $f(\xi) = \mu$. The theorem has thus been proved.

From equality (*) we find

$$\int_a^b f(x) dx = f(\xi)(b-a), \quad a < \xi < b$$

This formula enables us to restate the mean value theorem as follows:

The definite integral of a continuous function is equal to the product of the value of the function at some intermediate point of the interval of integration by the length of that interval.

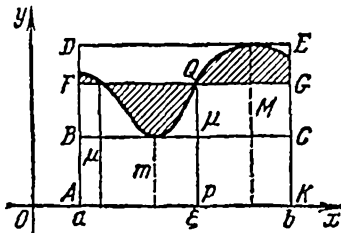


Fig. 109

Let us discuss a visual interpretation of this theorem. If a straight line parallel to Ox is moved upward (Fig. 109) from the position BC the area of the rectangle $ABCK$ increases continuously from a value less than the area of the trapezoid to a value exceeding it. Obviously, for some intermediate position of this line

(denote it by FG) the area of the rectangle $AFGK$ is precisely equal to the area of the trapezoid S . Since in this motion the straight line constantly intersects the curve bounding the trapezoid, in the position FG there appears one or several (two in Fig. 109) points of intersection Q , the abscissa of any such point of intersection being a value ξ indicated in the theorem.

IV. The arithmetic mean value of a function.

Definition. *The arithmetic mean value* (or, simply, *the mean*) y_m of a continuous function $y=f(x)$ on the interval $[a, b]$ is the ratio of the definite integral of this function to the length

of the interval:

$$y_m = \frac{\int_a^b f(x) dx}{b-a}$$

We shall present some considerations justifying this definition. Let a magnitude y take on n values $y_1, y_2, \dots, y_n$. The ratio $\frac{y_1 + y_2 + \dots + y_n}{n}$ is called the *arithmetic mean value* of this quantity. For example, if during the twenty-four hours the air temperature is measured every hour, the *mean temperature* is equal to the ratio of the sum of all the observed temperatures to 24.

Now suppose that a magnitude varies continuously (for instance, the air temperature is known at any moment of the 24 hours) and that we want to characterize "in the mean" the set of all its values. How should we define the mean air temperature taking into account the entire set of the data? More generally, what should be taken as the *mean value* of a continuous function $y = f(x)$ on an interval $[a, b]$?

To find the answer to the question, let us split the interval $[a, b]$ into n equal subintervals and take the values of the function at the midpoints of these subintervals serving as points ξ_i :

$$f(\xi_1), f(\xi_2), \dots, f(\xi_n)$$

Here the points ξ_i at which we take the values of the function are evenly spaced; this is a usual procedure in various kinds of measurement. Now we can take the arithmetic mean η_n of the chosen values:

$$\eta_n = \frac{f(\xi_1) + f(\xi_2) + \dots + f(\xi_n)}{n}$$

It is obvious that the greater n , that is, the greater the number of the values of the function which are taken into account in the determination of the mean value, the more reliable the result. Therefore it appears natural to take as the mean value y_m the *limit* to which η_n tends as $n \rightarrow \infty$. Let us find this limit.

On multiplying and dividing the expression of η_n by $b-a$ we obtain

$$\eta_n = \frac{1}{b-a} [f(\xi_1) \Delta x_1 + f(\xi_2) \Delta x_2 + \dots + f(\xi_n) \Delta x_n]$$

where $\Delta x_1 = \Delta x_2 = \dots = \Delta x_n = \frac{b-a}{n}$. Now, passing to the limit as $n \rightarrow \infty$ (i.e. as $\Delta x \rightarrow 0$) we get the expression given above for the mean value:

$$y_m = \lim_{n \rightarrow \infty} \frac{1}{b-a} \sum_{i=1}^n f(\xi_i) \Delta x_i = \frac{1}{b-a} \int_a^b f(x) dx$$

On the basis of the mean value theorem (Theorem VIII) we conclude that $y_{\overline{\eta}} = f(\xi)$ where $\xi \in (a, b)$. The mean value of a continuous function on a closed interval (provided the function is nonconstant) is always less than some of its values, greater than some other values and equal to at least one of the values.

The notion of the mean value of a function is commonly used in engineering. For instance, such quantities as vapour pressure, power of an alternating electric current, rate of a chemical reaction and the like are most often characterized by their mean values.

91. The Derivative of an Integral with Respect to its Upper Limit. Let us assume that the lower limit of the given integral is constant and the upper limit is variable. Making the upper limit to assume different values we obtain the corresponding values of the integral; consequently, under this condition, the integral becomes a *function* of its upper limit.

Let us discuss the question of notation. The independent variable in the upper limit is usually denoted by the same letter (say x) as the variable of integration. Thus, we write

$$I(x) = \int_a^x f(x) dx$$

However the letter x in the element of integration only serves to designate the *auxiliary variable* (the variable of integration) which runs over the values ranging from the lower limit a to the upper limit x as the integral is formed. If it is necessary to evaluate a particular value of the function $I(x)$, for instance, for $x=b$, i.e. $I(b)$, we substitute b for x in the upper limit of the integral but do not replace by b the variable of integration. Therefore it is more convenient to write

$$I(x) = \int_a^x f(t) dt$$

denoting the variable of integration by some other letter (in this case by t). However, for simplicity, we shall often denote by the same letter both the variable of integration and the independent variable in the upper limit bearing in mind that in the upper limit and under the integral sign they have different meanings.

The properties of the integral described in the previous section are also characteristic of an integral with variable upper limit.

For our further aims it is extremely important to study the connection between the function $I(x)$ and the given integrand $f(x)$.¹

¹ The presentation of the material in this course makes it possible to study the definite integral before the indefinite. It is only necessary to read Sec. 75 beforehand.

Theorem IX (On the Derivative of the Definite Integral with Respect to Its Upper Limit). The derivative of the integral with respect to its upper limit is equal to the integrand:

$$I'(x) = \left(\int_a^x f(x) dx \right)' = f(x) \quad (*)$$

In other words: an integral with variable upper limit is an anti-derivative of its integrand.

Proof. Let us give the argument x an increment Δx . Then the new value of the function $I(x)$ is

$$I(x + \Delta x) = \int_a^{x+\Delta x} f(x) dx$$

Hence,

$$\Delta I = I(x + \Delta x) - I(x) = \int_a^{x+\Delta x} f(x) dx - \int_a^x f(x) dx$$

On applying Theorem IV to the first integral on the right, we get

$$\Delta I = \int_a^x f(x) dx + \int_x^{x+\Delta x} f(x) dx - \int_a^x f(x) dx = \int_x^{x+\Delta x} f(x) dx$$

According to the mean value theorem, the last integral is representable as

$$\Delta I = f(\xi)(x + \Delta x - x) = f(\xi)\Delta x$$

where ξ is a point lying between x and $x + \Delta x$.

By the definition of the derivative, we have

$$I'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta I}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{f(\xi)\Delta x}{\Delta x} = \lim_{\Delta x \rightarrow 0} f(\xi)$$

But if $\Delta x \rightarrow 0$ then $x + \Delta x$ tends to x , and therefore $\xi \rightarrow x$, which implies, since $f(x)$ is a continuous function, that

$$\lim_{\Delta x \rightarrow 0} f(\xi) = \lim_{\xi \rightarrow x} f(\xi) = f(x)$$

This is what we wanted to prove.

From this theorem it follows that

$$d \int_a^x f(x) dx = f(x) dx \quad (**)$$

It should be noted that the results expressed by formulas (*) and (**) do not depend on the notation of the variable of integ-

ration; for instance, we can also write

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

and

$$d \int_a^x f(t) dt = f(x) dx$$

The geometrical interpretation of the theorem was indicated in Sec. 78; here we shall explain it once more. The function $I(x)$ expresses the *variable* area of the curvilinear trapezoid with variable base $[a, x]$ bounded by the curve $y = f(x)$. The assertion of the theorem implies that the derivative of the area of the trapezoid with respect to the abscissa is equal to the ordinate of the line bounding the trapezoid (i.e. to the segment $AB = f(x)$; see Fig. 110) or that the differential of the area of the trapezoid is equal to the area of the rectangle $ABDE$ with sides equal, res-

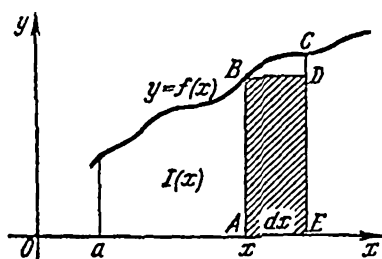


Fig. 110

pectively, to the increment of the base of the trapezoid and to the ordinate of the line $y = f(x)$ at the point x .

92. The Newton-Leibniz Theorem. Now we have come to the concluding stage of our considerations which enables us to establish a roundabout way for evaluating integrals without computing the limits of integral sums.

Theorem X (The Newton-Leibniz Theorem). The value of the definite integral is equal to the difference of the values of any antiderivative of the integrand corresponding to the upper and lower limits of the integral:

$$\int_a^b f(x) dx = F(b) - F(a), \text{ where } F'(x) = f(x) \quad (*)$$

The equality $(*)$ is referred to as the Newton-Leibniz formula.

In other words:

The value of the definite integral is equal to the increment of any antiderivative of the integrand gained on the interval of integration.

Proof. Let us consider the function

$$I(x) = \int_a^x f(x) dx$$

Since it is an antiderivative of the function $f(x)$, it must be contained, according to the theorem of Sec. 78, in the family of functions $F(x) + C$ where $F(x)$ is an *arbitrary* antiderivative of $f(x)$. Consequently,

$$I(x) = F(x) + C_1$$

where C_1 is a concrete constant. To determine C_1 we shall make use of another property of the function $I(x)$, namely,

$$I(a) = \int_a^a f(x) dx = 0$$

It now follows that

$$F(a) + C_1 = 0, \text{ i.e. } C_1 = -F(a)$$

Thus,

$$I(x) = \int_a^x f(x) dx = F(x) - F(a)$$

Finally, putting $x=b$ we obtain the desired equality (*).

The difference of the values of a function is often written in the symbolic form:

$$F(b) - F(a) = F(x) \Big|_a^b$$

The symbol $\Big|_a^b$ with the two indices a and b is the so-called *sign of double substitution*. It indicates that the value of the function corresponding to the lower index must be subtracted from the one corresponding to the upper index.

Using this notation, we can write formula (*) in the form

$$\int_a^b f(x) dx = F(x) \Big|_a^b$$

where

$$F'(x) = f(x)$$

The Newton-Leibniz theorem provides a key method for evaluating definite integrals. It enables us to find the definite integral with the aid of antiderivatives, i.e. with the aid of *indefinite integration*, avoiding summation. To illustrate this method let us take some simple examples:

$$\int_1^2 x^2 dx = \frac{x^3}{3} \Big|_1^2 = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}$$

$$\int_a^b \frac{1}{x} dx = \ln x \Big|_a^b = \ln b - \ln a = \ln \frac{b}{a} \quad (a > 0, b > 0)$$

$$\int_0^{\frac{\pi}{2}} \cos x dx = \sin x \Big|_0^{\frac{\pi}{2}} = 1$$

It is natural to call the complex function $Z(t)$ an *antiderivative* of the complex function $z(t)$. The indefinite integral of $z(t)$ is the family of all its antiderivatives; it can be readily shown that

$$\int z(t) dt = Z(t) + C$$

where $C = C_1 + iC_2$ is an arbitrary complex constant (C_1 and C_2 are real constants). We can also write

$$\int z(t) dt = \int x(t) dt + i \int y(t) dt$$

The definite integral of the function $z(t)$ is defined similarly:

$$\int_a^b z(t) dt = \int_a^b x(t) dt + i \int_a^b y(t) dt$$

On applying the Newton-Leibniz formula to the integrals on the right-hand side of the equality we obtain

$$\int_a^b z(t) dt = X(t) \Big|_a^b + iY(t) \Big|_a^b = Z(t) \Big|_a^b$$

This is the *Newton-Leibniz formula for complex functions of a real argument*.

Examples. (1) If $n \neq 0$ is an integer, then

$$\int_0^{2\pi} e^{int} dt = \frac{e^{int}}{in} \Big|_0^{2\pi} = \frac{e^{2\pi in} - 1}{in} = 0$$

because $e^{2\pi in} = 1$ (see Sec. 74).

$$(2) \text{ We have } \int_0^{\pi} e^{it} dt = \frac{e^{it}}{i} \Big|_0^{\pi} = \frac{e^{i\pi} - 1}{i} = -\frac{2}{i} = 2i$$

(3) Let $I_1 = \int e^{at} \cos bt dt$ and $I_2 = \int e^{at} \sin bt dt$. The method of finding these integrals by means of integrating by parts is described in Sec. 81, I (Example 4). The integrals I_1 and I_2 can be found more easily by means of Euler's formula (see Sec. 74):

$$I_1 + iI_2 = \int e^{at} (\cos bt + i \sin bt) dt = \int e^{(a+ib)t} dt = \frac{e^{(a+ib)t}}{a+ib} + C$$

where C is a complex constant. Separating the real and imaginary parts on the right-hand side of the equality we obtain the expressions for the sought-for integrals. Let the reader carry out the calculations; the final results are given in the table of integrals at the end of the book (Formulas 73 and 74).

§ 3. Methods of Evaluating Definite Integrals

94. **Integration by Parts and by Change of Variable in the Definite Integral.** The most important result of the theory of integration is the Newton-Leibniz formula which establishes the connection between the definite and the indefinite integrals; according to this formula the definite integral is evaluated with the aid of the indefinite integral:

$$\int_a^b f(x) dx = \int f(x) dx \Big|_a^b$$

Since we can take any antiderivative on the right-hand side of the equality, it makes no difference what arbitrary constant is taken in the indefinite integration and therefore it is usually omitted.

The simplest rules for integrating a sum of functions and a product of a function by a constant are essentially the same for the definite and the indefinite integrals. Now we proceed to the rules of integration by parts and by change of variable and indicate some particulars related to their application for evaluating definite integrals.

I. **Integration by parts:** we have

$$\int_{x_1}^{x_2} u dv = uv \Big|_{x_1}^{x_2} = \int_{x_1}^{x_2} v du \quad (A)$$

where u and v are functions of an independent variable.

Proof. The relation

$$\int_{x_1}^{x_2} u dv = \int_{x_1}^{x_2} u dv \Big|_{x_1}^{x_2} = \left(uv - \int v du \right) \Big|_{x_1}^{x_2}$$

directly implies the formula we set out to prove.

Instead of completing indefinite integration by parts and performing double substitution we can make use of formula (A).

As an example, let us take the integral

$$I_n = \int_0^{\frac{\pi}{2}} \sin^n x dx \text{ where } n \text{ is a positive integer.}$$

We put

$$dv = \sin x dx, \quad u = \sin^{n-1} x$$

Then

$$v = -\cos x, \quad d\mu = (n-1) \sin^{n-2} x \cos x dx$$

and

$$I_n = -\cos x \sin^{n-1} x \Big|_0^{\frac{\pi}{2}} + (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x \cos^2 x dx$$

The first summand on the right-hand side is equal to zero; replacing $\cos^2 x$ by $1 - \sin^2 x$ in the second summand we obtain

$$I_n = (n-1) \int_0^{\frac{\pi}{2}} \sin^{n-2} x dx - (n-1) \int_0^{\frac{\pi}{2}} \sin^n x dx$$

i.e.

$$I_n = (n-1) I_{n-2} - (n-1) I_n$$

whence

$$I_n = \frac{n-1}{n} I_{n-2}$$

This is the *recurrence formula* for the integrals I_n . With the aid of this formula we can finally reduce the exponent to 1 if n is an odd number and to zero if n is an even number. The integrals

$$I_1 = \int_0^{\frac{\pi}{2}} \sin x dx = 1, \quad I_0 = \int_0^{\frac{\pi}{2}} dx = \frac{\pi}{2}$$

being known, the problem can be thus completely solved. For example,

$$I_3 = \int_0^{\frac{\pi}{2}} \sin^3 x dx = \frac{2}{3} I_1 = \frac{2}{3}, \quad I_4 = \int_0^{\frac{\pi}{2}} \sin^4 x dx = \frac{3}{4} I_2 = \frac{3}{4} \cdot \frac{1}{2} I_0 = \frac{3}{16} \pi$$

The determination of the indefinite integral corresponding to the original definite integral would involve more complicated calculations.

II. Change of variable in the definite integral (integration by substitution). The rule for change of variable in the definite integral is very important. We remind the reader (see Sec. 81, II) that if $x = \psi(u)$ then

$$\int f(x) dx = \int f[\psi(u)] \psi'(u) du$$

In this case if one of the integrals is known, the other is found as well. If the first integral is equal to $F(x)$ (the arbitrary constant C is omitted), the second is equal to $F[\psi(u)]$; if the second

integral is known, in order to find the first integral we must replace u by its expression in terms of x .

We can now state the rule:

Let $f(x)$ be a continuous function in a closed interval $[x_1, x_2]$, and let $x = \psi(u)$ and $\psi'(u)$ be continuous functions in a closed interval $[u_1, u_2]$, the function $\psi(u)$ being monotone in this interval, and $\psi(u_1) = x_1$, $\psi(u_2) = x_2$.¹ Then

$$\int_{x_1}^{x_2} f(x) dx = \int_{u_1}^{u_2} f[\psi(u)] \psi'(u) du \quad (B)$$

Proof. We assume that the indefinite integral on the left-hand side is known and is equal to $F(x)$; then

$$\int_{x_1}^{x_2} f(x) dx = F(x_2) - F(x_1)$$

According to the above, the indefinite integral on the right-hand side is equal to $F[\psi(u)]$ and, therefore,

$$\int_{u_1}^{u_2} f[\psi(u)] \psi'(u) du = F[\psi(u_2)] - F[\psi(u_1)] = F(x_2) - F(x_1)$$

Comparing the equalities, we arrive at Formula (B).

Formula (B) shows that the element of integration is transformed as in the case of the indefinite integral. The new limits of integration u_1 and u_2 are the roots of the equations

$$x_1 = \psi(u) \text{ and } x_2 = \psi(u)$$

in the unknown u .

Thus, instead of performing indefinite integration with the aid of change of variable, returning to the original variable and then evaluating the double substitution with the *given* limits, we can directly compute the double substitution with the *new* limits. The result (the value of the definite integral) is the same but the amount of calculations is reduced.

The change of variable in the definite integral is often performed not by the formula $x = \psi(u)$ but by the formula $u = \varphi(x)$ expressing the new variable in terms of the old one. Then the new limits u_1 and u_2 are readily found by the formulas

$$u_1 = \varphi(x_1), \quad u_2 = \varphi(x_2)$$

¹ Under these conditions the composite function $f[\psi(u)]$ is defined and continuous in the interval $[u_1, u_2]$; by the way, usually it is possible to discard the condition of monotonicity of the function $\psi(x)$.

Sometimes there may arise difficulties caused by the fact that if the function $u = \varphi(x)$ is nonmonotone, its inverse function $x = \psi(u)$ is not single-valued. It may turn out that to different values x_1 and x_2 there correspond the same values $u_1 = u_2$ (for instance, this is the case in the double substitution for $u = \cos x$ with the limits $-\frac{\pi}{2}$ and $\frac{\pi}{2}$) and the integral on the right-hand side of formula (B) is equal to zero whereas that on the left is not. Therefore, for formula (B) to be valid some additional conditions should be imposed on the function $u = \varphi(x)$; for instance, it is sufficient to require that the function $\varphi(x)$ *should be monotone in the interval $[x_1, x_2]$ and have a derivative not vanishing at any interior point of the interval* (in short, the derivative $\varphi'(x)$ should retain its sign). Then the inverse function $x = \psi(u)$ satisfies all the requirements stated in the rule for integration by substitution (see Theorem V in Sec. 34 and also Sec. 42, II).

Examples. (1) Consider the integral $\int_0^{\frac{1}{2}} x^2 \sqrt{1-x^2} dx$. We put $x = \sin u$ and find the new limits of integration u_1 and u_2 from the equations $0 = \sin u$ and $\frac{1}{2} = \sin u$; u_1 can be taken equal to 0, and u_2 equal to $\frac{\pi}{6}$. As u varies from 0 to $\frac{\pi}{6}$ the variable $x = \sin u$ runs throughout the given interval of integration $\left[0, \frac{1}{2}\right]$. Thus,

$$\begin{aligned} \int_0^{\frac{1}{2}} x^2 \sqrt{1-x^2} dx &= \int_0^{\frac{\pi}{6}} \sin^2 u \cos^2 u du = \frac{1}{4} \int_0^{\frac{\pi}{6}} \sin^2 2u du = \\ &= \frac{1}{8} \left(u - \frac{\sin 4u}{4} \right) \Big|_0^{\frac{\pi}{6}} = \frac{\pi}{48} - \frac{\sqrt{3}}{64} \end{aligned}$$

(2) Let us find $\int_0^{\frac{\pi}{2}} \frac{\sin x}{1 + \cos^2 x} dx$. The substitution $u = \cos x$ satisfies all the requirements in the interval $0 \leq x \leq \frac{\pi}{2}$; therefore,

$$\int_0^{\frac{\pi}{2}} \frac{\sin x}{1 + \cos^2 x} dx = - \int_1^0 \frac{du}{1+u^2} = \arctan u \Big|_0^1 = \frac{\pi}{4}$$

(3) Taking the integral $U_n = \int_0^{\frac{\pi}{2}} \cos^n x dx$ we put $x = \frac{\pi}{2} - u$; then

$$\int_0^{\frac{\pi}{2}} \cos^n x dx = - \int_{\frac{\pi}{2}}^0 \cos^n \left(\frac{\pi}{2} - u \right) du = \int_0^{\frac{\pi}{2}} \sin^n u du$$

and since the value of the definite integral does not depend on the notation of the variable of integration we obtain

$$U_n = \int_0^{\frac{\pi}{2}} \cos^n x dx = \int_0^{\frac{\pi}{2}} \sin^n x dx$$

The method for evaluating the last integral with the aid of the recurrence formulas was described on page 319.

Let us derive some formulas for an integral

$$\int_{-a}^a f(x) dx$$

taken over a symmetric interval $[-a, a]$ whose integrand is an even or an odd function. We first represent this integral as

$$\int_{-a}^a f(x) dx = \int_{-a}^0 f(x) dx + \int_0^a f(x) dx$$

Then, changing the variable of integration in the first integral on the right-hand side by means of the formula $x = -u$, we obtain

$$\int_{-a}^a f(x) dx = \int_0^a f(-u) du + \int_0^a f(x) dx$$

In the first integral on the right we designate the variable of integration by x and derive the formula

$$\int_{-a}^a f(x) dx = \int_0^a [f(-x) + f(x)] dx$$

The integrand on the right-hand side is equal to zero if $f(x)$ is an odd function and to $2f(x)$ if $f(x)$ is an even function. Consequently,

$$\int_{-a}^a f(x) dx = \begin{cases} 0 & \text{for an odd function } f(x) \\ 2 \int_0^a f(x) dx & \text{for an even function } f(x) \end{cases}$$

These formulas are very useful. For instance, it immediately becomes clear that

$$\int_{-a}^a x^5 e^{x^2} dx = 0 \quad \text{and} \quad \int_{-\pi}^{\pi} \sin^3 x \cos^2 x dx = 0$$

We advise the reader to elucidate the geometrical meaning of these formulas.

95. Approximate Integration. We now turn to some commonly used methods of *approximate integration* enabling us to find with a sufficient accuracy an approximate value of the definite integral of any continuous function. The need for an approximate evaluation of the integral may arise when the method for finding the exact value of the integral does not exist or is unknown or when the application of the existing method is inconvenient. The approximate numerical methods considered here are based on the following:

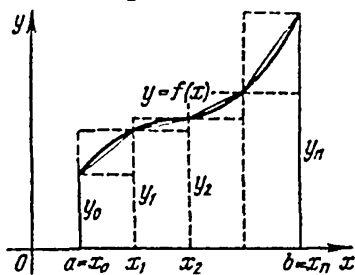


Fig. 111

regarding the integral as the area of the corresponding curvilinear trapezoid we can find an approximate value of that area and thus an approximate value of the integral if we compute the area of another trapezoid bounded by a curve whose shape and position are as close as possible to those of the given curve. The auxiliary curve should be chosen so that the area of the new figure is easily computed.

We shall consider the following rules for numerical integration: (1) the *rectangle rule* and the *trapezoid rule* and (2) *Simpson's rule*¹.

I. The rectangle rule and the trapezoid rule. Let us divide the given interval of integration $[a, b]$ into n equal parts (subintervals) and replace the given trapezoid by a step-like figure consisting of n rectangles with these subintervals as bases, the altitudes of the rectangles being equal to the values of the function $y = f(x)$ at the initial or final points of the subintervals (Fig. 111). The value of the area of such a figure gives an approximation to the

sought-for integral $I = \int_a^b f(x) dx$. The greater the number of the subintervals, the greater the accuracy of the result.

If we denote the values of the function $f(x)$ at the points of division by $y_0, y_1, \dots, y_{n-1}, y_n$, i.e. if we put $y_k = f(x_k)$ and

¹ T. Simpson (1710-1761), an English mathematician.

$x_k = a + k\Delta x$ where $\Delta x = \frac{b-a}{n}$ and k takes on the values $0, 1, 2, \dots, n$, we obviously get the following approximate formulas:

$$I \approx \underline{I} = \Delta x (y_0 + y_1 + \dots + y_{n-1})$$

and

$$I \approx \bar{I} = \Delta x (y_1 + y_2 + \dots + y_n)$$

These are the so-called *rectangle formulas*.

Now let us retain the same division of the interval $[a, b]$ but replace the arcs of the curve $y = f(x)$ corresponding to the subintervals by the chords joining the end points of these arcs. The curvilinear trapezoid is then replaced by a figure consisting of n rectilinear trapezoids (Fig. 111). As a rule, the area of such a figure expresses the sought-for area more accurately than the area of the step figure made up of rectangles.

It is obvious that the area of the trapezoid constructed on a subinterval is equal to half the sum of the areas of the rectangles corresponding to this subinterval. Adding up all these areas we obtain

$$I \approx \frac{\underline{I} + \bar{I}}{2} = \Delta x \left(\frac{y_0 + y_n}{2} + y_1 + y_2 + \dots + y_{n-1} \right)$$

This relation is known as the *trapezoid formula*.

If the function $f(x)$ is monotone, $\underline{I}$ and $\bar{I}$ are bounds for the integral providing a considerably more accurate estimation than that given in Theorem VI, Sec. 90. It is clear that if the function is *monotone increasing* (which is the case in Fig. 111) we have

$$\underline{I} < I < \bar{I}$$

and if it is *monotone decreasing*, then, conversely,

$$\bar{I} > I > \underline{I}$$

In these cases, taking as an approximation to the integral the value obtained by the trapezoid formula

$$I \approx \frac{\underline{I} + \bar{I}}{2}$$

we get the result with an absolute error not exceeding the absolute value of half the difference between $\bar{I}$ and $\underline{I}$, i.e.

$$|\Delta I| \leq \frac{|\bar{I} - \underline{I}|}{2}$$

As an example, let us find approximate values of the integral

$$\int_0^1 \frac{dx}{1+x^2}$$

and compare them with its exact value equal to $\arctan 1 = \frac{\pi}{4} \approx 0.785398$. Let $n=10$. Then $\Delta x=0.1$ and $x_k=k \cdot 0.1$ ($k=0, 1, 2, \dots, 10$). Computing the values of the function (to within 0.001) we get

$$y_0=1, \quad y_1=\frac{1}{1+(0.1)^2} \approx 0.990,$$

$$y_2=\frac{1}{1+(0.2)^2} \approx 0.962, \quad y_3 \approx 0.917, \quad y_4 \approx 0.862, \quad y_5 \approx 0.8,$$

$$y_6 \approx 0.735, \quad y_7 \approx 0.671, \quad y_8 \approx 0.610, \quad y_9 \approx 0.552, \quad y_{10}=0.5$$

Consequently,

$$\begin{aligned} \underline{I} = 0.1 (1 + 0.990 + 0.962 + 0.917 + 0.862 + 0.8 + 0.735 + \\ + 0.671 + 0.610 + 0.552) \approx 0.810 \end{aligned}$$

We similarly compute $\bar{I}$:

$$\bar{I} \approx 0.760$$

In this example $\underline{I} > \bar{I}$ because the function is decreasing. Thus,

$$0.760 < I < 0.810$$

As an approximation to I we now take the value

$$I \approx \frac{0.760 + 0.810}{2} \approx 0.785$$

Here the absolute error does not exceed $|\Delta I| < 0.025$, the relative error being $\frac{0.025}{0.785} \cdot 100 \approx 3.2\%$. A more thorough analysis shows that the absolute error is in fact considerably smaller (it does not exceed 0.0004).

Here we shall also compute an approximate value of the integral

$$I = \int_0^{\pi} \sin x \, dx = 2$$

Let us put $n=6$. Then, by the trapezoid formula,

$$\begin{aligned} I &\approx \frac{\pi}{6} \left(\frac{\sin 0 + \sin \pi}{2} + \sin \frac{\pi}{6} + \sin \frac{2\pi}{6} + \sin \frac{3\pi}{6} + \sin \frac{4\pi}{6} + \sin \frac{5\pi}{6} \right) = \\ &= \frac{\pi}{6} \left(0.5 + \frac{1}{2}\sqrt{3} + 1 + \frac{1}{2}\sqrt{3} + 0.5 \right) \approx 1.9541 \end{aligned}$$

The absolute error is equal to 0.0459 and the relative error is 2.5%.

In this example the integrand is not monotone in the interval of integration, and therefore it is difficult to estimate the integral. Indeed, here we have $\underline{I} = \bar{I}$, i.e. we only obtain a minor approximation of the sought-for integral. In order to find an estimation for this integral we must break up the interval of integration into parts in which the function is monotone and investigate separately the integrals over these parts.

II. Simpson's rule. This rule does not require much more labour than the previous ones but yields more accurate results for the same divisions of the interval of integration.

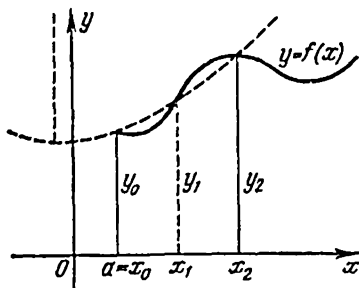


Fig. 112

As before, we divide the interval $[a, b]$ into n equal parts but this time we assume that n is even: $n = 2m$. The arc of the curve $y = f(x)$ corresponding to the interval $[x_0, x_2]$ is now replaced by the parabola through the three points (x_0, y_0) , (x_1, y_1) and (x_2, y_2) with axis parallel to Oy (Fig. 112) where x_0 , x_1 , and x_2 are, respectively, the initial point, the centre point

and the terminal point of the subinterval $[x_0, x_2]$. Analytically this means that in the interval $[x_0, x_2]$ the given function $y = f(x)$ is replaced by a quadratic function $y = px^2 + qx + r$. The coefficients p , q and r should be chosen so that the values of both functions are equal for x_0 , x_1 and x_2 :

$$y_0 = px_0^2 + qx_0 + r, \quad y_1 = px_1^2 + qx_1 + r, \quad y_2 = px_2^2 + qx_2 + r$$

On solving these equations we find the coefficients p , q and r .

Making similar transformations in the intervals $[x_2, x_4]$, $[x_4, x_6]$, $\dots$, $[x_{n-2}, x_n]$, we assume that the area of the given trapezoid is approximately equal to the sum of the areas of the *parabolic trapezoids* we have obtained. Our next task is to prove that the area S of the trapezoid bounded by a parabola $y = px^2 + qx + r$ with axis parallel to the axis of ordinates is expressed by the formula

$$S = \frac{(\text{base length})}{6} (y_{in} + 4y_c + y_t) \quad (*)$$

where y_{in} , y_c and y_t are, respectively, the ordinates of the parabola at the initial, centre and terminal points of the base of the trapezoid.

Let us first suppose that the base of the trapezoid is an interval $[-\gamma, \gamma]$ of the x -axis symmetric with respect to the origin. For

the area of such a parabolic trapezoid we have

$$S = \int_{-\gamma}^{\gamma} (px^2 + qx + r) dx = p \int_{-\gamma}^{\gamma} x^2 dx + q \int_{-\gamma}^{\gamma} x dx + r \int_{-\gamma}^{\gamma} dx$$

i.e.

$$S = p \frac{\gamma^3 - (-\gamma)^3}{3} + q \frac{\gamma^2 - (-\gamma)^2}{2} + r \frac{\gamma - (-\gamma)}{1} = \frac{2}{3} p\gamma^3 + 2r\gamma$$

Since here we have $y_{in} = y_{x=-\gamma} = p\gamma^2 - q\gamma + r$, $y_c = y_{x=0} = r$, $y_t = y_{x=\gamma} = p\gamma^2 + q\gamma + r$ and the base length is equal to 2γ , the direct substitution of these values into formula (*) shows us that it holds.

It is obvious that this formula remains valid for a parabolic trapezoid of the type under consideration with any base. Indeed, the area of the trapezoid is not changed under a parallel displacement carrying the centre point of its base to the origin: after the displacement the area can be computed by formula (*).

Let us now return to the original problem and apply the formula obtained to find the area S_1 of the parabolic trapezoid with the interval $[x_0, x_1]$ as base:

$$S_1 = \frac{x_1 - x_0}{6} (y_0 + 4y_1 + y_2) = \frac{\Delta x}{3} (y_0 + 4y_1 + y_2)$$

where $\Delta x = \frac{b-a}{n}$. The areas $S_2, S_3, \dots, S_m$ of the subsequent parabolic trapezoids are similarly expressed as

$$\begin{aligned} S_2 &= \frac{\Delta x}{3} (y_2 + 4y_3 + y_4) \\ S_3 &= \frac{\Delta x}{3} (y_4 + 4y_5 + y_6) \\ &\dots \dots \dots \\ S_m &= \frac{\Delta x}{3} (y_{n-2} + 4y_{n-1} + y_n) \end{aligned}$$

Adding up all these equalities termwise we obtain the expression giving an approximate value of the sought-for integral:

$$I \approx \frac{\Delta x}{3} [y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \dots + 4y_{n-1} + y_n]$$

We have thus obtained *Simpson's formula*.

For instance, by Simpson's formula, taking $n=10$, we find for the integral $I = \int_0^1 x^4 dx$ (whose exact value is 0.2) the approximate value 0.200013; the absolute error is 0.000013 and the relative error is 0.01%.

It is clear that for any integral of a quadratic function $y = px^2 + qx + r$ Simpson's formula gives the exact value. It should also be noted that for an integral of a cubic function Simpson's formula gives the exact value as well.

Let us apply Simpson's formula to the integral $\int_0^{\pi} \sin x dx$ evaluated above by means of the trapezoid formula.

For $n=6$ simple computations yield

$$I \approx \frac{\pi}{18} (8 + 2\sqrt{3}) \approx 2.001$$

The relative error of this result is equal to 0.05%. The accuracy is high although there are only six points of division of the interval of integration.

Applying to the integral $\int_0^1 \frac{dx}{1+x^2}$ Simpson's formula with $n=4$ we find $I \approx 0.78539$. Since $I = \frac{\pi}{4}$ we obtain $\pi = 3.14156$. The absolute error of this approximate value of the number π is less than $4 \cdot 10^{-5}$. Earlier, when applying the trapezoid rule, we obtained for the value of this integral (and, hence, for the number π) a less accurate result although the working was more laborious.

Still more accurate formulas of approximate integration are treated in special courses devoted to numerical methods of analysis (for instance, see [7], [12], [13]).

The above rules for approximate numerical integration enable us to find integrals not only of functions represented by formulas but also of those specified by graphs or by tables.

Example. The width of a river is 20 m; the measurements of the depth in its cross-section made with a step of 2 m resulted in the following table:

x	0	2	4	6	8	10	12	14	16	18	20
y	0.2	0.5	0.9	1.1	1.3	1.7	2.1	1.5	1.1	0.6	0.2

Here x is the distance (in metres) reckoned from one of the banks and y the corresponding depth (also in metres). It is required to determine the area S of the cross-section of the river.

The *trapezoid formula* yields

$$S = 2 \left(\frac{0.2 + 0.2}{2} + 0.5 + 0.9 + 1.1 + 1.3 + 1.7 + \right. \\ \left. + 2.1 + 1.5 + 1.1 + 0.6 \right) = 22 \text{ m}^2$$

and Simpson's formula gives

$$S = \frac{2}{3} [0.2 + 4 \cdot 0.5 + 2 \cdot 0.9 + \dots + 4 \cdot 0.6 + 0.2] = 21.9 \text{ m}^2$$

The results are very close. There is no sense in discussing their accuracy since the conditions of the problem give no exact information about the profile of the cross-section of the river.

96. Graphical Integration. Now we shall discuss a graphical method of evaluating definite integrals. The problem is posed as follows: given the graph of a function $y=f(x)$, it is required to find, purely geometrically, without resorting to calculations, an approximate value of the integral

$$I = \int_a^b f(x) dx$$

which is equal to the area of the curvilinear trapezoid bounded by the line $y=f(x)$.

Let us choose equal scales along Ox and Oy . Mark on the x -axis the point P lying at unit distance to the left of the origin (Fig. 113), and draw a straight line FG parallel to the x -axis intersecting the curve $y=f(x)$ so that the area of the rectangle $AFGB$ differs as little as possible from that of the curvilinear trapezoid. This means that the area enclosed between the curve and the straight line FG and lying above the line should be as close as possible to the area enclosed between the curve and that straight line and lying below the line. Let us continue the straight line FG to intersect Oy at a point T and connect this point with P by a line segment. Finally, from the point A we draw a straight line parallel to PT which intersects the straight line $x=b$ at a point M' .

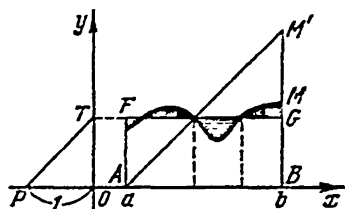


Fig. 113

It can readily be shown that the segment $M'B$ represents the sought-for area, i.e. the number of units of length contained in the segment $M'B$ is equal to the number of units of area (corresponding to the given scale) in the area of the curvilinear trapezoid. Indeed, the similarity of the triangles POT and ABM' implies

$$\frac{BM'}{AB} = \frac{OT}{OP} \text{ whence } BM' = \frac{AB \cdot OT}{OP}$$

where $OP=1$ and $AB \cdot OT$ is equal to the area of the trapezoid.

If the interval $[a, b]$ is not sufficiently small the error introduced by an inaccurate construction of the straight line FG may be

substantial. For a more accurate construction the interval of integration is divided into subintervals (not necessarily equal) and the whole trapezoid into the corresponding number of trapezoids with these subintervals as bases. The points of division should be chosen so that each subinterval is an interval of monotonicity of the integrand and every point of intersection of the curve $y=f(x)$ with the x -axis is among the points of division. If these subintervals are sufficiently small so that in each of them the curve only

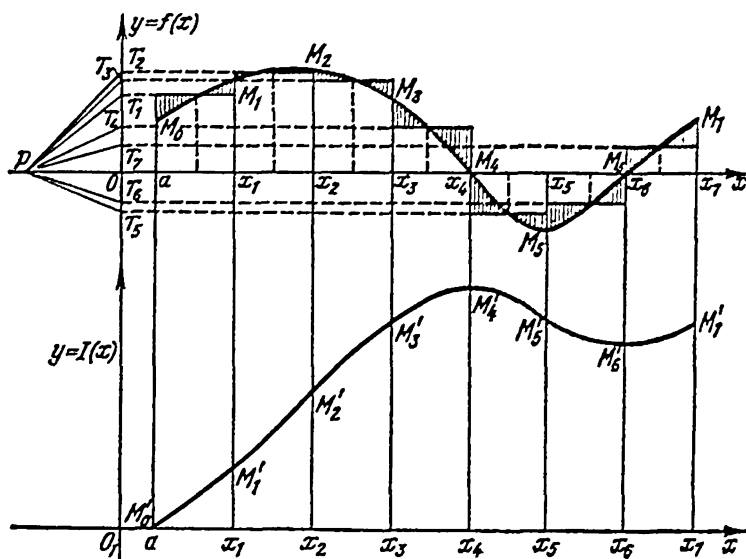


Fig. 114

slightly differs from a straight line we can take, as midlines of the subtrapezoids, the ordinates of the graph of $y=f(x)$ at the midpoints of the subintervals. This technique is usually applied in practice since it makes it unnecessary to draw by eye the lines of the type of FG .

Thus, we construct, in succession, for each of the curvilinear subtrapezoids, a line segment representing its area. It is more convenient to lay off these segments from another axis O_1x parallel to Ox (Fig. 114). The area reckoned from the straight line $x=a$, corresponding to the point $x=a$ is obviously equal to zero. We mark on O_1x the point $M'_0(a, 0)$ which corresponds to the point M_0 of the curve $y=f(x)$. To the point $x=x_1$ there corresponds the area of the first subtrapezoid; it is represented by the segment M'_0x_1 , which is obtained if we draw from the point M'_0 a straight line parallel to PT_1 intersecting the line $x=x_1$ at M'_1 . This point corresponds to the point M_1 of the curve. The area of the trapezoid corresponding to the straight line $x=x_2$ (i.e. the value of the

integral from a to x_2) is equal to the sum (understood algebraically) of the areas of the first and the second subtrapezoids. It is represented by the segment M_2x_2 , which appears as we draw from the point M_1' a straight line parallel to PT_1 , intersecting the straight line $x=x_2$ at M_2' . The point M_2' corresponds to the point M_2 of the curve. Proceeding in this manner we construct, in succession, the points $M_3', M_4', \dots$, corresponding to the points $M_3, M_4, \dots$ of the curve. The ordinate of the point M_n' corresponding to the point $M_n[b, f(b)]$ of the curve gives us the required approximate value of the integral I . Obviously, the greater the number of points of division, the more accurate the approximation.

Let us join the points obtained $M_0', M_1', M_2', \dots, M_n'$ by a continuous curve. The ordinates of this curve obviously represent (approximately) the values of the integral taken from $x=a$ to the corresponding points of the base of the trapezoid. In other words, this curve is the graph of the function

$$I(x) = \int_a^x f(x) dx$$

determined by the integral with variable upper limit.

As was said (Sec. 79), the curve $y=I(x)$ is called an *integral curve* of the function $y=f(x)$. The geometrical construction of the integral curve with the aid of the graph of the integrand presented here is called *graphical integration*.

As in the case of graphical differentiation, graphical integration is most convenient when the integrand is represented graphically and its analytical expression is unknown. Such a situation often occurs in practice; for example, this is the case when a function is specified by a graph traced by a self-recording apparatus.

In conclusion we note that for approximate integration of functions represented graphically, i.e. for approximate computation of the areas of the corresponding curvilinear trapezoids, there are special instruments called *planimeters*. After the tracer of the planimeter has been passed around the periphery of the area the value of the area is read on the meter of the planimeter. There also exist instruments (called *integragraphs*) for constructing an approximate integral curve of a given function from its graph.

§ 4. Improper Integrals

97. Infinite Integrals. Up to now, when speaking of definite integrals we assumed that the interval of integration was finite and closed and that the integrand was *continuous*. It is this particular case to which the existence theorem for the definite integral stated in Sec. 87 applies. However, it often becomes necessary to

extend the definition of the definite integral to an infinite interval of integration or to an unbounded integrand function. In this section we shall consider the case of infinite limits (the so-called *infinite integrals*); integrals of discontinuous functions will be studied in Sec. 99.

Let $y=f(x)$ be a continuous function in an infinite interval $[a, +\infty)$, i.e. for $x \geq a$ (in this case it is allowable, for brevity, to omit the plus sign in front of the symbol ∞ and write $[a, \infty)$). Then we can take the integral of the function $f(x)$ over any finite interval $[a, l]$ ($l > a$):

$$I(l) = \int_a^l f(x) dx$$

Now let us make l grow indefinitely. Then there are two possibilities, namely, $I(l)$ either has a limit as $l \rightarrow +\infty$ or has no limit, which justifies the following definition:

Definition. The *improper integral* $\int_a^\infty f(x) dx$ of the function $f(x)$ over the interval $[a, \infty)$ is the limit of the integral $\int_a^l f(x) dx$ as $l \rightarrow \infty$ provided that this limit exists:

$$\int_a^\infty f(x) dx = \lim_{l \rightarrow \infty} \int_a^l f(x) dx \quad (*)$$

If the limit exists the improper integral $\int_a^\infty f(x) dx$ is said to be *convergent*. If the limit does not exist equality (*) is meaningless, and the improper integral $\int_a^\infty f(x) dx$ is said to be *divergent*. In this case $\int_a^l f(x) dx$ either tends to infinity or has no limit at all.

If an antiderivative $F(x)$ of the integrand $f(x)$ is known, it can easily be checked whether the given improper integral is convergent or divergent: if $\lim_{l \rightarrow \infty} F(l) = F(\infty)$ exists, then, by the Newton-Leibniz formula,

$$\int_a^\infty f(x) dx = \lim_{l \rightarrow \infty} [F(l) - F(a)] = F(\infty) - F(a)$$

and the integral is convergent; if this limit does not exist the integral is divergent.

Examples. (1) $\int_0^{\infty} e^{-x} dx = -e^{-x} \Big|_0^{\infty} = 0 + 1 = 1$. Consequently, the improper integral $\int_0^{\infty} e^{-x} dx$ is convergent and equal to 1.

(2) $\int_1^{\infty} \frac{dx}{x} = \ln x \Big|_1^{\infty} = \infty$. The integral is divergent because the antiderivative $\ln x$ tends to infinity as $x \rightarrow \infty$.

(3) The integral $\int_0^{\infty} \cos x dx$ is divergent since the integral $\int_0^l \cos x dx = \sin x \Big|_0^l = \sin l$ has no limit as $l \rightarrow \infty$ (it oscillates between -1 and 1).

The improper integral over the interval $(-\infty, a]$ is defined similarly

$$\int_{-\infty}^a f(x) dx = \lim_{l \rightarrow -\infty} \int_l^a f(x) dx = F(a) - F(-\infty)$$

where $F(-\infty)$ is the limit (if it exists) of the antiderivative $F(x)$ as $x \rightarrow -\infty$.

If $f(x)$ is a function continuous throughout Ox the improper integral can be defined for the whole interval $(-\infty, +\infty)$: by definition,

$$\int_{-\infty}^{+\infty} f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^{+\infty} f(x) dx$$

If both integrals on the right-hand side are convergent, the integral $\int_{-\infty}^{+\infty} f(x) dx$ is said to be convergent. If at least one of the integrals on the right-hand side is divergent the equality has no sense, and the integral on the left is called divergent.

If an antiderivative $F(x)$ is known, then

$$\int_{-\infty}^{+\infty} f(x) dx = F(+\infty) - F(-\infty)$$

where $F(+\infty)$ and $F(-\infty)$ are, respectively, the limits (if they exist) of $F(x)$ as $x \rightarrow +\infty$ and $x \rightarrow -\infty$. If at least one of these

limits does not exist the improper integral is divergent

$$(4) \int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \arctan x \Big|_{-\infty}^{+\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi$$

$$(5) \int_{-\infty}^{+\infty} e^x dx = e^x \Big|_{-\infty}^{+\infty} = \infty - 0 = \infty. \text{ The integral is divergent.}$$

$$(6) \int_{-\infty}^{+\infty} \frac{x}{1+x^2} dx = \frac{1}{2} \ln(1+x^2) \Big|_{-\infty}^{+\infty}. \text{ Here both } F(-\infty) \text{ and}$$

$F(+\infty)$ are equal to infinity, and the integral is divergent.

It should be noted that all the simplest properties of definite integrals enumerated in Sec. 88 are extended without any changes to improper integrals provided all the integrals on the right-hand sides of the equalities printed in medium-faced type are convergent. Theorems III-V of Sec. 89 also remain true for improper integrals.

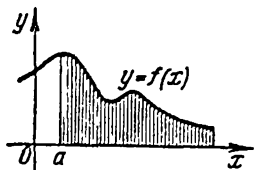


Fig. 115

A convergent improper integral can be interpreted geometrically. For instance, suppose that the graph of a function $y = f(x)$ bounds a trapezoid with infinite base (Fig. 115). If the improper integral $\int_a^\infty f(x) dx$ is convergent we say that the area of the shaded figure is equal to this integral. If the integral is divergent it is impossible to speak of the area of that figure.

For instance, to the infinite trapezoid whose boundary consists of the positive half-axis Ox , the straight line $x = a$ ($a > 0$) and the curve $y = \frac{1}{x^3}$ we can attribute an area equal to $\frac{1}{2a^2}$ because

$$\int_a^\infty \frac{1}{x^3} dx = -\frac{1}{2} \frac{1}{x^2} \Big|_a^\infty = \frac{1}{2a^2}$$

But with the infinite trapezoid bounded by the hyperbola $y = \frac{1}{x}$, the positive half-axis Ox and the straight line $x = a$ ($a > 0$) we cannot associate an area because

$$\int_a^\infty \frac{1}{x} dx = \ln x \Big|_a^\infty = \infty$$

Improper integrals of the type considered in this section are often encountered in problems of mechanics and electrostatics in connection with the notion of a *potential*. Let a point M of mass

m placed at the origin attract a moving point M_1 of unit mass lying on Ox at a distance of x from M . As is known, by Newton's law of gravitation, the magnitude P of the force of attraction is specified by the formula

$$P = k \frac{m}{x^2}$$

where k is a constant, while the work performed as M_1 is moved from the point $x=r$ to the point $x=b$ ($b > r > 0$) is expressed by the formula

$$A = - \int_r^b k \frac{m}{x^2} dx = km \left(\frac{1}{b} - \frac{1}{r} \right)$$

The minus sign in front of the integral is accounted for by the fact that the direction of the force is opposite to that of the motion of the point M_1 (by the same reason, the work is negative).

If the point M_1 recedes to infinity ($b = \infty$), then

$$A = - \int_r^{\infty} k \frac{m}{x^2} dx = -k \frac{m}{r}$$

In the case when the point M_1 moves from infinity to the point $x=r$ the force of attraction performs a *positive* work:

$$A = k \frac{m}{r}$$

The magnitude of this work is called *the potential of the force of attraction of the material point M at the point $x=r$* .

98*. Tests for Convergence of Improper Integrals with Infinite Limits. The investigation of the convergence of an improper integral becomes substantially complicated if the corresponding antiderivative is unknown. In such a case it is sometimes possible to find out whether the integral is convergent or divergent with the aid of some special *tests* without resorting to the antiderivative.

For definiteness, let us speak about integrals of the form $\int_a^{\infty} f(x) dx$;

the tests for convergence of integrals $\int_{-\infty}^a f(x) dx$ are completely analogous.

I. Let us begin with the case when the integrand $f(x)$ is nonnegative at all the points of the interval $[a, \infty)$: $f(x) \geq 0$. It is clear, geometrically, that $I(l) = \int_a^l f(x) dx$ increases together with l

(if there are intervals where $f(x) \equiv 0$, the integral $I(l)$ is constant on them). According to the test for existence of a limit given at the end of Sec. 31, if an increasing function is *bounded* it possesses a limit. Therefore, if it is known that $I(l)$ is bounded, that is if $I(l) \leq M$ where M is a constant, we can assert that the

improper integral $\int_a^\infty f(x) dx$ is convergent and that its value does not exceed M . At the same time, it is apparent that these considerations provide no means for practical evaluation of the integral. When the antiderivative is unknown the computation of the improper integral is much more complicated than the investigation of its convergence and requires some special techniques.

Proceeding from what has been said we can readily prove the following test:

Comparison Test. Let the inequality

$$0 \leq f(x) \leq \varphi(x) \quad (*)$$

hold for all the values of x . Then: (1) if the integral $\int_a^\infty \varphi(x) dx$

is convergent the integral $\int_a^\infty f(x) dx$ is convergent as well;

(2) If the integral $\int_a^\infty f(x) dx$ is divergent the integral $\int_a^\infty \varphi(x) dx$ is also divergent.

Proof. Suppose that the integral $\int_a^\infty \varphi(x) dx$ is convergent and is equal to M ; then $\int_a^l \varphi(x) dx \leq M$ for any l . By the theorem on integrating inequalities (Sec. 90, II), we have

$$\int_a^l f(x) dx \leq \int_a^l \varphi(x) dx \leq M$$

This means that $I(l) = \int_a^l f(x) dx$ is a bounded function, and therefore the integral $\int_a^\infty f(x) dx$ is convergent.

Now, let it be known that the integral $\int_a^\infty f(x) dx$ is divergent;

then the increasing function $\int_a^l f(x) dx$ tends to $+\infty$. But since $\int_a^l \varphi(x) dx \geq \int_a^l f(x) dx$, the function $\int_a^l \varphi(x) dx$ also tends to $+\infty$, i.e. the integral $\int_a^\infty \varphi(x) dx$ is divergent. The proof of the test has thus been completed.

Let the reader elucidate the geometrical meaning of the comparison test.

Instead of requiring that inequality (*) should hold for the whole interval $[a, \infty)$ we can suppose that it is fulfilled for all $x \geq b > a$. In this case the test also remains valid since we can apply it to the integral over the interval $[b, \infty)$ and then add to that integral the integral from a to b . The latter being a constant quantity, this obviously does not violate the convergence or the divergence.

The main difficulty in investigating the convergence of a given improper integral $\int_a^\infty f(x) dx$ by means of the comparison test is that it is not known in advance with what function the integrand should be compared. Therefore, to apply the test we must have at our disposal a collection of functions whose integrals are known to be convergent or divergent. For instance, it can readily be verified that the integral of the function $y = e^{-kx}$ over any interval $[a, \infty)$ is convergent for $k > 0$ and divergent for $k \leq 0$.

Since for $x > 1$ we have the inequality $e^{-x^2} < e^{-x}$ the integral $\int_1^\infty e^{-x^2} dx$ is convergent. By the above remark, the integral $\int_0^\infty e^{-x^2} dx$ is also convergent, and, since the integrand is even, the integral $\int_{-\infty}^\infty e^{-x^2} dx$ is convergent as well.

The integral $\int_{-\infty}^{+\infty} e^{-x^2} dx$ is known as the *Euler-Poisson¹ integral*; it plays an important role in the *probability theory*. Later on (Sec. 131) we shall evaluate this integral by means of a special technique; here we simply write down its numerical value:

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$$

¹ Poisson, Simeon Denis (1781-1840), a French mathematician and physicist.

As a function with which the integrand of a given integral is compared the power function $y = \frac{1}{x^m}$ is frequently used. For this function we shall prove the following assertion:

The improper integral $\int_a^\infty \frac{dx}{x^m}$ where $a > 0$ is convergent for $m > 1$ and divergent for $m \leq 1$. (It is obvious that only the case $m > 0$ is of interest because for $m < 0$ the integrand tends to infinity as $x \rightarrow \infty$ and therefore the integral of that function is apparently divergent.)

Indeed, for $m \neq 1$ we have

$$\int_a^\infty \frac{dx}{x^m} = \frac{x^{-m+1}}{-m+1} \Big|_a^\infty$$

If $m > 1$ the exponent $1-m$ is negative, and the antiderivative tends to zero as $x \rightarrow \infty$; therefore the integral is convergent. If $m < 1$, then $1-m > 0$, and the antiderivative tends to infinity as $x \rightarrow \infty$, i.e. the integral is divergent. The divergence of the integral for $m=1$ was established in Example 2 of Sec. 97.

Examples. (1) The integral $\int_1^\infty \frac{dx}{\sqrt{1+x} \sqrt[3]{1+x^2}}$ is convergent because the inequality $\frac{1}{\sqrt{1+x} \sqrt[3]{1+x^2}} < \frac{1}{x^{1/2} x^{2/3}} = \frac{1}{x^{7/6}}$ holds for all $x \geq 1$ and the integral $\int_1^\infty \frac{dx}{x^{7/6}}$ is convergent since $m = \frac{7}{6} > 1$.

(2) The integral $\int_1^\infty \frac{\sqrt{x}}{1+x} dx$ is divergent since $\frac{\sqrt{x}}{1+x} > \frac{\sqrt{x}}{x+x} = \frac{1}{2\sqrt{x}}$ for $x > 1$ and the integral $\int_1^\infty \frac{dx}{\sqrt{x}}$ is divergent because $m = \frac{1}{2} < 1$.

II. The comparison test stated in I only applies to functions retaining constant sign within the infinite interval of integration (it is clear that the investigation of the integral of a negative function reduces to the foregoing case). It is more difficult to investigate integrals of functions with alternating sign, for instance, such as $\frac{\sin x}{x}$. Here we give a convergence test for such integrals

which sometimes makes it possible to reduce their investigation to the case of a positive integrand.

If the integral $\int_a^\infty |f(x)| dx$ of the absolute value of the function $f(x)$ is convergent the integral $\int_a^\infty f(x) dx$ is also convergent. In this case the integral $\int_a^\infty f(x) dx$ is said to be *absolutely convergent*.

To prove the test we take two auxiliary functions $f^+(x)$ and $f^-(x)$ determined by the relations

$$f^+(x) = \begin{cases} f(x) & \text{for } f(x) \geq 0 \\ 0 & \text{for } f(x) < 0 \end{cases}$$

and

$$f^-(x) = \begin{cases} 0 & \text{for } f(x) > 0 \\ f(x) & \text{for } f(x) \leq 0 \end{cases}$$

(The construction of the graphs of the functions $f^+(x)$ and $f^-(x)$ for a given function $f(x)$ is shown in Fig. 116.) It is obvious that the functions $f^+(x)$ and $-f^-(x)$ are *nonnegative*, and each of them

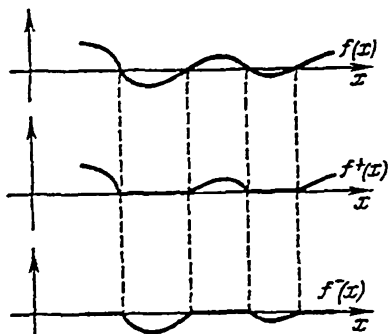


Fig. 116

does not exceed $|f(x)|$. Since, according to the hypothesis, the integral of $|f(x)|$ is convergent, we conclude, by the comparison test, that the integrals $\int_a^\infty f^+(x) dx$ and $\int_a^\infty [-f^-(x)] dx$ are convergent as well; the change of the sign of the integrand in the last integral does not violate its convergence.

But $f(x) = f^+(x) + f^-(x)$, and, since the integrals of $f^+(x)$ and $f^-(x)$ are convergent, the integral $\int_a^\infty f(x) dx$ is also convergent (see the remark on page 334). The test has thus been proved.

For example, the integrals $\int_0^\infty \frac{\cos x}{1+x^2} dx$ and $\int_0^\infty \frac{\sin x}{1+x^2} dx$ are absolutely convergent because the absolute values of the integrands do not exceed the positive function $\frac{1}{1+x^2}$, the integral of the latter being convergent (see Example 4 in Sec. 97).

If the integral $\int_{-\infty}^{+\infty} |f(x)| dx$ is convergent, the integrals $\int_{-\infty}^{+\infty} f(x) \cos x dx$ and $\int_{-\infty}^{+\infty} f(x) \sin x dx$ are absolutely convergent since the absolute values of the integrands obviously do not exceed $|f(x)|$. In what follows we shall take advantage of this simple fact.

The fact that the integral of $|f(x)|$ is divergent does not enable us to judge upon the convergence of the integral of $f(x)$ because it can be either divergent or convergent. In the latter case (i. e. when the integral of $f(x)$ is convergent and the integral of $|f(x)|$ is divergent) we say that the integral $\int_a^{\infty} f(x) dx$ is *conditionally* (not absolutely) *convergent*.

An example of this kind is the *Dirichlet*¹ integral $\int_0^{\infty} \frac{\sin x}{x} dx$. It can be shown (we do not present the proof here) that this integral is convergent while the integral of the absolute value of the integrand is divergent. Thus, the Dirichlet integral is conditionally convergent. Its numerical value is found by means of some special techniques:

$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$$

99. Integrals of Discontinuous Functions. If a function $f(x)$ has a number of points of discontinuity of the first kind in an interval $[a, b]$ the integral of this function is defined as the sum of the ordinary integrals taken over the subintervals into which the interval $[a, b]$ is broken up by all the points of discontinuity of the function. Denoting these points as $c_1, c_2, \dots, c_k$ ($a < c_1 < c_2 < \dots < c_k < b$) we write

$$\int_a^b f(x) dx = \int_a^{c_1} f(x) dx + \int_{c_1}^{c_2} f(x) dx + \dots + \int_{c_k}^b f(x) dx$$

In the first integral on the right-hand side the value of the function $f(x)$ at the point c_1 is understood as its *left-hand limit* $f(c_1 - 0)$, and in the second integral as the *right-hand limit* $f(c_1 + 0)$ (see Sec. 33). The values of the function $f(x)$ at the other points of discontinuity are understood similarly. Under the conditions assumed, in every closed subinterval of integration the integrand is

¹ Dirichlet, Peter Gustav Lejeune (1805-1859), a German mathematician.

continuous. The geometrical meaning of the integral under consideration is clear from Fig. 117: the integral is equal to the sum of the areas of the trapezoids having as bases the subintervals $[a, c_1]$, $[c_1, c_2]$, ..., $[c_k, b]$ lying between the subsequent points of discontinuity.

We now proceed to the definition of the improper integral for functions with *infinite discontinuities*. Let $y=f(x)$ be a continuous function for all $x \in [a, b)$ (i. e., $a \leq x < b$) having an infinite discontinuity at the right end point $x=b$ of the interval $[a, b]$. It is clear that the ordinary definition of the definite integral is inapplicable here. Let us first take the ordinary integral

$$I(\varepsilon) = \int_a^{b-\varepsilon} f(x) dx \text{ where } \varepsilon > 0$$

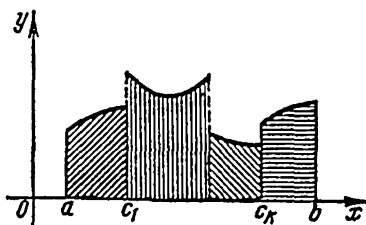


Fig. 117

and then make ε tend to zero.

Then $I(\varepsilon)$ either tends to a finite limit or has no finite limit (in the latter case it either tends to infinity or has no limit at all).

Definition. The *improper integral* $\int_a^b f(x) dx$ of a function $f(x)$ continuous for $a \leq x < b$ and unbounded as $x \rightarrow b$ is the limit of the integral $\int_a^{b-\varepsilon} f(x) dx$ for $\varepsilon \rightarrow 0$ ($\varepsilon > 0$) provided this limit exists:

$$\int_a^b f(x) dx = \lim_{\varepsilon \rightarrow 0} \int_a^{b-\varepsilon} f(x) dx, \varepsilon > 0$$

If the limit exists we say that the improper integral is *convergent*. If this limit does not exist the equality becomes meaningless, and the improper integral written on the left is said to be *divergent*.

Similarly, if the function $f(x)$ has an infinite discontinuity only at the left end point $x=a$ of the interval $[a, b]$ we put

$$\int_a^b f(x) dx = \lim_{\delta \rightarrow 0} \int_{a+\delta}^b f(x) dx \text{ where } \delta > 0$$

provided this limit exists.

If the antiderivative $F(x)$ is known we can write, in both cases,

$$\int_a^b f(x) dx = F(b) - F(a)$$

where $F(b)$ ($F(a)$) is understood as the limit of the antiderivative $F(x)$ as $x \rightarrow b$ (as $x \rightarrow a$) under the assumption that this limit exists. If the limit does not exist, the integral is divergent.

Examples. (1) $\int_0^a \frac{dx}{\sqrt{x}} = 2\sqrt{x} \Big|_0^a = 2\sqrt{a}.$

(2) $\int_0^1 \frac{dx}{1-x} = -\ln(1-x) \Big|_0^1.$ For $x \rightarrow 1$ the limit of $\ln(1-x)$ is equal to infinity; therefore the integral is divergent.

Generally, observe that the integrals $\int_a^b \frac{dx}{(x-a)^k}$ and $\int_a^b \frac{dx}{(b-x)^k}$ are convergent if $k < 1$ and divergent if $k \geq 1$. (The proof is left to the reader.)

(3) $\int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} = \arcsin x \Big|_{-1}^1 = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi.$ In this example the

integrand has infinite discontinuities at both end points of the interval of integration.

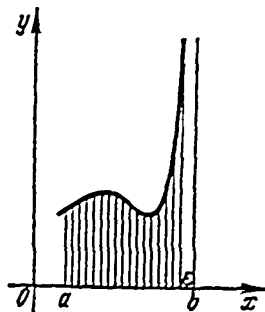


Fig. 118

If the function $f(x)$ has an infinite discontinuity at an intermediate point $x=c$ of the interval $[a, b]$ (i. e. $a < c < b$) then, by definition,

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

If both integrals on the right-hand side of the equality are convergent the integral

$\int_a^b f(x) dx$ is also convergent; if at least one of these integrals on the right is divergent the integral is divergent.

(4) $\int_{-1}^2 \frac{dx}{\sqrt[3]{x^2}} = \int_{-1}^0 \frac{dx}{\sqrt[3]{x^2}} + \int_0^2 \frac{dx}{\sqrt[3]{x^2}} = 3\sqrt[3]{x} \Big|_{-1}^0 + 3\sqrt[3]{x} \Big|_0^2 = 3 + 3\sqrt[3]{2}.$

(5) $\int_{-1}^1 \frac{dx}{x^2} = \int_{-1}^0 \frac{dx}{x^2} + \int_0^1 \frac{dx}{x^2}.$ Both integrals on the right-hand side

are divergent, and, consequently, the given integral is also divergent.

Suppose that the curve $y=f(x)$ has an asymptote perpendicular to Ox at the point $x=b$; then the trapezoid bounded by it is infinite (with an "infinite altitude"; see Fig. 118). If the improper

integral of the function $f(x)$ exists the area of the infinite trapezoid is, by definition, equal to the value of this integral; if otherwise, we do not assign any area to this trapezoid.

For instance, to the infinite trapezoid bounded by the curve $y = \frac{1}{\sqrt{x}}$ and the straight lines $x=0$, $x=a$ and $y=0$ we can at-

tribute an area equal to $2\sqrt{a}$ since $\int_0^a \frac{dx}{\sqrt{x}} = 2\sqrt{a}$ (see Example 1).

At the same time, with the infinite trapezoid bounded by the hyperbola $y = \frac{1}{x}$ and the same straight lines $x=0$, $x=a$ and $y=0$

we can associate no area since the integral $\int_0^a \frac{dx}{x} = \ln x \Big|_0^a$ is divergent.

Here we do not dwell on tests for convergence of integrals of functions with infinite discontinuities since they are analogous to those discussed in Sec. 98.

QUESTIONS

1. What is an antiderivative of a given function? Give examples.
2. State the theorem on the antiderivatives of a continuous function. What is the general form of the antiderivative of a given continuous function?
3. What is the indefinite integral of a given function?
4. State and prove the simplest properties of the indefinite integral. Give examples of their application to indefinite integration.
5. Describe the methods of change of variable and of integration by parts for the indefinite integral. Demonstrate them by examples.
6. What rational fraction is called proper? What fractions are called partial rational fractions?
7. How is the decomposition of a proper rational fraction into partial fractions performed?
8. How do we integrate rational functions?
9. Give examples of integration of the simplest irrational functions.
10. What is the general method for finding an integral of a function rational with respect to trigonometric functions?
11. How do we compute the integrals of the form $\int \sin^n x \cos^m x dx$ where n and m are integers?
12. When do we say that a function is nonintegrable in elementary functions (in finite form)?

13. How do we define the area of a curvilinear trapezoid, the work of a force, the path length travelled and the mass?

14. What is the definite integral of a given function over a given interval? Express the notions enumerated in Question 13 in terms of the integral.

15. State the theorem on the existence of the definite integral.

16. State and prove the simplest properties of the definite integral.

17. Explain the properties of additivity and of retention of sign of the definite integral.

18. What is the geometrical meaning of the definite integral of a given function $y=f(x)$ over an interval $[a, b]$ in the Cartesian coordinates?

19. State and prove the theorem on the estimation of the integral and give its geometrical interpretation. Prove the generalized theorem on the estimation of the integral.

20. State and prove the mean value theorem for the definite integral and illustrate it geometrically.

21. What is the arithmetical mean of a function $y=f(x)$ on an interval $[a, b]$?

22. What is the derivative of the integral with respect to the upper limit of integration equal to? Prove the theorem on the derivative and give its geometrical interpretation.

23. State and derive the Newton-Leibniz rule.

24. How is the method of integration by parts applied to the definite integral?

25. How is the method of change of variable used in the definite integration?

26. Derive the formulas simplifying the expression of the definite integral of an even and of an odd function over a symmetric interval $[-a, a]$.

27. State and derive the rules for an approximate computation of the definite integral: (1) the rectangle and the trapezoid rules, (2) Simpson's rule.

28. Describe the method of graphical integration.

29. State the definition of the improper integral of a given function over an infinite interval. Illustrate this notion geometrically and give examples of convergent and divergent integrals.

30. Prove the comparison test for improper integrals.

31. What is an absolutely convergent improper integral? What improper integral is said to be conditionally convergent?

32. State the definition of the improper integral of a function with infinite discontinuity over a given finite interval. Illustrate this notion geometrically and give examples of convergent and divergent integrals of that kind.

Chapter VI

APPLICATION OF INTEGRAL CALCULUS

§ 1. Some Problems of Geometry and Statics

100. Area of a Plane Figure. We begin with computing the area or, as we say, the *quadrature* or *squaring* of a plane figure. Historically, this problem is so closely related to integration that integrals themselves are often called *quadratures* and a problem reducible to evaluating an integral is said to be *solvable (integrable) by quadratures*.

1. Area in Cartesian coordinates. In Cartesian coordinates a curvilinear trapezoid is the simplest (basic) figure whose area is expressed by one integral. If $y=f(x)$ is the equation of the line bounding the trapezoid, and $y \geq 0$, the area S of the trapezoid is given by the formula

$$S = \int_a^b y dx$$

where the limits of integration a and b ($a < b$) are the abscissas of the initial and the terminal points of the line.

If the line $y=f(x)$ is specified by parametric equations

$$x = \varphi(t), \quad y = \psi(t)$$

the substitution $x = \varphi(t)$ brings the integral expressing the area to the form

$$S = \int_{t_1}^{t_2} \psi(t) \varphi'(t) dt$$

where t_1 and t_2 are the limits between which the parameter t varies as the variable point runs from left to right over the upper boundary of the trapezoid (i.e. the line $y=f(x)$).

Example. Let us find the area S bounded by the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. We have

$$S = 2 \int_{-a}^a y dx, \quad y \geq 0$$

Here it is convenient to pass to the parametric representation $x = a \cos t$, $y = b \sin t$ (see Sec. 46), which yields

$$S = -2ab \int_{\pi}^0 \sin^2 t \, dt$$

The computation results in

$$S = 2ab \int_0^{\pi} \sin^2 t \, dt = 2ab \frac{1}{2} \left(t - \frac{\sin 2t}{2} \right) \Big|_0^{\pi} = \pi ab$$

II. Area in polar coordinates. If the line bounding a figure is represented by an equation in polar coordinates, then, as the basic figure, we take a *curvilinear sector* (Fig. 119) which is an area bounded by a line $r = f(\varphi)$ such that every ray issued from the pole P cuts it at no more than one point and by the two rays $\varphi = \alpha$ and $\varphi = \beta$ (α and β are measured in radians).

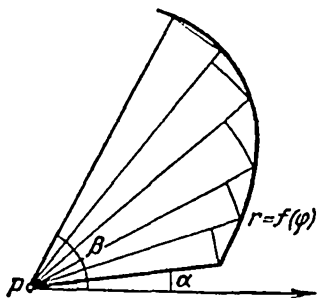


Fig. 119

Let us show that the area of such a sector can be expressed by one integral. We break up the entire sector into n subsectors with the help of the rays whose angles of inclination to the polar axis are

$$\varphi_0 = \alpha, \varphi_1, \varphi_2, \dots, \varphi_n = \beta$$

The next step is to replace every *curvilinear* subsector by a *circular* one, that is, the function $f(\varphi)$ is considered, approximately, to be constant and equal to $r_k = f(\varphi_k)$ on each subinterval $[\varphi_k, \varphi_{k+1}]$ ($k = 0, 1, \dots, n-1$).

We remind the reader that the area of a circular sector of radius r with central angle α is equal to $\frac{1}{2} r^2 \alpha$. Therefore the area of the figure made up of n circular sectors replacing the curvilinear sectors (Fig. 119) is

$$\begin{aligned} S_n &= \frac{1}{2} f^2(\varphi_0) (\varphi_1 - \varphi_0) + \frac{1}{2} f^2(\varphi_1) (\varphi_2 - \varphi_1) + \dots \\ &\dots + \frac{1}{2} f^2(\varphi_{n-1}) (\varphi_n - \varphi_{n-1}) = \frac{1}{2} \sum_{i=0}^{n-1} f^2(\varphi_i) \Delta\varphi_i \end{aligned} \quad (*)$$

where $\Delta\varphi_i = \varphi_{i+1} - \varphi_i$.

For simplicity, when forming integral sum (*) we choose the intermediate points ξ_i coinciding with the *left* end points of the subintervals $[\varphi_i, \varphi_{i+1}]$. This results in an inessential change in the notation of $\Delta\varphi_i$, and the summation, with respect to i , is carried out from 0 to $n-1$ (instead of from 1 to n).

If $n \rightarrow \infty$ and the greatest of the angles $\Delta\varphi_i$ tends to zero the limit of sum (*) is exactly the sought-for area of the curvilinear sector. Expression (*) is an integral sum for the function $f^2(\varphi)$ on the interval $[\alpha, \beta]$, and therefore

$$S = \frac{1}{2} \int_{\alpha}^{\beta} f^2(\varphi) d\varphi$$

or, briefly,

$$S = \frac{1}{2} \int_{\alpha}^{\beta} r^2 d\varphi \quad \text{where } r = f(\varphi)$$

Examples. (1) Let us compute the area S of the figure bounded by the first winding of the spiral of Archimedes $r = a\varphi$ and by the corresponding segment of the polar axis (Fig. 120):

$$S = \frac{a^2}{2} \int_0^{2\pi} \varphi^2 d\varphi = \frac{4}{3} \pi^3 a^2$$

(2) Let us compute the area S of the figure bounded by the curve $r = a \cos 3\varphi$ (referred to as the *three-leafed rose*; see Fig. 121). The

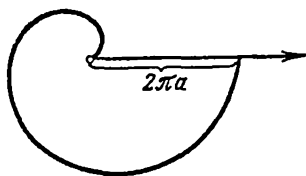


Fig. 120

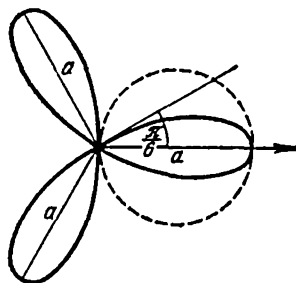


Fig. 121

area of one sixth of the whole figure (i.e. half the area of one loop) is expressed by the integral

$$\frac{1}{2} a^2 \int_0^{\frac{\pi}{6}} \cos^2 3\varphi d\varphi$$

For, the terminus of the polar radius traverses the whole line bounding this sector as the polar angle varies from 0 to $\frac{\pi}{6}$.

Consequently, the area of the rose is

$$S = 3a^2 \int_0^{\frac{\pi}{6}} \cos^2 3\varphi \, d\varphi = a^2 \int_0^{\frac{\pi}{2}} \cos^2 \psi \, d\psi = \frac{\pi a^2}{4}$$

(we have made the change $3\varphi = \psi$); thus, it is equal to the area of a circle with diameter a .

101. Volume of a Solid¹. Let there be given a body bounded by a closed surface and let the *area of its any cross-section* by a plane perpendicular to a fixed straight line be known; for instance, this line may be the axis of the abscissas (see Fig. 122).

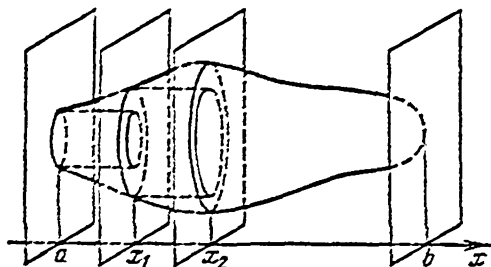


Fig. 122

We shall assume that the area of the cross-section is a continuous function $S(x)$ where x is the abscissa of the point of intersection of the cutting plane with the x -axis.

Furthermore, suppose that the whole solid is contained between two planes perpendicular to the x -axis and cutting it at points a and b ($a < b$). To evaluate the volume of such a solid we divide it (mentally) into layers with the aid of cutting planes perpendicular to the x -axis and intersecting it at points $x_0=a, x_1, x_2, \dots, x_n=b$ and replace each layer by a *right cylinder* of the same altitude with base $S(x_i)$ (see Fig. 122). The volume of a right cylinder is equal to the product of the area of its base by the altitude, and therefore, the volume of the n -step body is expressed by the sum

$$\begin{aligned} V_n &= S(x_0)(x_1 - x_0) + S(x_1)(x_2 - x_1) + \dots + S(x_{n-1})(x_n - x_{n-1}) = \\ &= \sum_{i=0}^{n-1} S(x_i) \Delta x_i \end{aligned}$$

The limit of this sum (which is an integral sum for the function $S(x)$ on the interval $[a, b]$, as $n \rightarrow \infty$ and $\max \Delta x_i \rightarrow 0$) is the desired volume:

$$V = \int_a^b S(x) \, dx \quad (*)$$

¹ The general definition of the volume of a space figure and its computation are connected with double integrals and will be given in Chapter VIII. Here we confine ourselves to an important particular case of the problem.

Fig. 122 shows a body of a simple shape for which the right cylinders thus constructed are either entirely contained within the corresponding layers or, conversely, envelop these layers. The resultant formula also remains valid when this condition does not hold provided that the contours bounding cross-sections lying close to each other do not "differ considerably" (we do not go into particulars connected with this observation).

If the body under consideration is obtained by the rotation of a curvilinear trapezoid bounded by the line $y = f(x)$ about Ox the cross-section with abscissa x is a circle with radius equal to the corresponding ordinate of the line $y = f(x)$. (If $y < 0$ the radius is equal to $|y|$.) In this case

$$S(x) = \pi y^2$$

and we thus arrive at the formula for the volume of a solid of revolution:

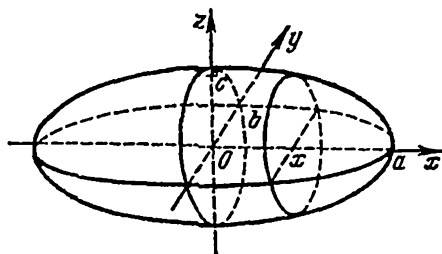


Fig. 123

$$V = \pi \int_a^b y^2 dx \quad \text{where } y = f(x)$$

Examples. (1) Let us find the volume V of the triaxial ellipsoid¹

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

Its plane sections perpendicular, for instance, to Ox (see Fig. 123) are ellipses with semi-axes $b \sqrt{1 - \frac{x^2}{a^2}}$ and $c \sqrt{1 - \frac{x^2}{a^2}}$ ($-a \leq x \leq a$).

The cross-section area $S(x)$ at the point x is known (see Sec. 100):

$$S(x) = \pi b \sqrt{1 - \frac{x^2}{a^2}} c \sqrt{1 - \frac{x^2}{a^2}} = \pi bc \left(1 - \frac{x^2}{a^2}\right)$$

Consequently,

$$V = \pi bc \int_{-a}^a \left(1 - \frac{x^2}{a^2}\right) dx = 2\pi bc \int_0^a \left(1 - \frac{x^2}{a^2}\right) dx = 2\pi bc \left(x - \frac{x^3}{3a^2}\right) \Big|_0^a$$

i.e.

$$V = \frac{4}{3} \pi abc$$

¹ If the reader is not yet familiar with algebraic surfaces of the second order this example can be omitted.

If there are two equal semi-axes, for instance, $c=b$, we have an ellipsoid of revolution with volume $V = \frac{4}{3}\pi ab^2$. If all the three semi-axes are equal, i.e. $a=b=c$, the ellipsoid turns into a sphere

of volume $V = \frac{4}{3}\pi a^3$.

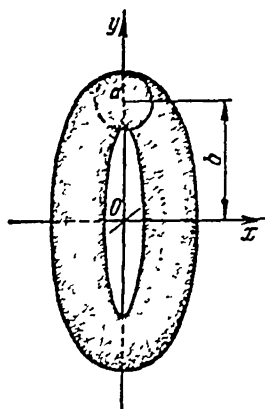


Fig. 124

(2) Let us compute the volume V of a *torus* (anchor ring) which is a body generated by the rotation in space of a circle about an axis in its plane but not cutting the circle. Let a circle of radius a rotate about Ox , its centre (in the initial position) being at a point $(0, b)$ $b > a$ (Fig. 124). The equation of the circle is

$$x^2 + (y-b)^2 = a^2$$

Apparently, the desired volume is found by subtracting the volume of the body generated by the rotation of the lower semicircle with the equation $y = b - \sqrt{a^2 - x^2}$ from the volume of the body generated by the rotation of the upper semicircle ($y = b + \sqrt{a^2 - x^2}$). Hence, we have

$$\begin{aligned} V &= \pi \int_{-a}^a (b + \sqrt{a^2 - x^2})^2 dx - \pi \int_{-a}^a (b - \sqrt{a^2 - x^2})^2 dx = \\ &= 2\pi \int_0^a [(b + \sqrt{a^2 - x^2})^2 - (b - \sqrt{a^2 - x^2})^2] dx = 8\pi b \int_0^a \sqrt{a^2 - x^2} dx \end{aligned}$$

Finally, putting $x = a \sin t$ and evaluating the integral we obtain

$$V = 8\pi b a^2 \int_0^{\frac{\pi}{2}} \cos^2 t dt = 8\pi b a^2 \frac{\pi}{4} = 2\pi^2 a^2 b$$

102. Arc Length. Let a curve be the graph of a continuous function $y = f(x)$ whose derivative $f'(x)$ is also continuous. As was said in Sec. 66, such a curve is said to be *smooth*, and its *arc length* is defined as follows:

The arc length of a curve is the limit of the length of the polygonal line inscribed into the curve as the number of its segments increases indefinitely and the length of the greatest segment tends to zero.

Thus, let a curve AB be represented by the equation

$$y = f(x)$$

We shall suppose that x varies within a closed interval where the function $f(x)$ and its first derivative $f'(x)$ are continuous. Let us divide the arc AB into n parts, the points of division having abscissas $x_0 = a, x_1, x_2, \dots, b = x_n$ (Fig. 125). Drawing a chord through every pair of subsequent points of division we construct an inscribed broken line whose length L_n is expressed by the formula

$$\begin{aligned} L_n &= \sqrt{\Delta x_1^2 + \Delta y_1^2} + \\ &+ \sqrt{\Delta x_2^2 + \Delta y_2^2} + \dots + \sqrt{\Delta x_n^2 + \Delta y_n^2} = \\ &= \sum_{i=1}^n \sqrt{\Delta x_i^2 + \Delta y_i^2} \end{aligned}$$

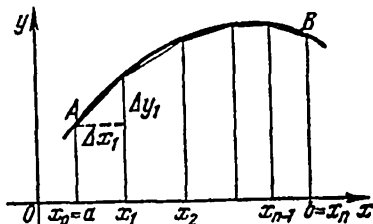


Fig. 125

that is,

$$L = \sum_{i=1}^n \sqrt{1 + \left(\frac{\Delta y_i}{\Delta x_i}\right)^2} \Delta x_i$$

where $\Delta x_i = x_i - x_{i-1}$ and $\Delta y_i = f(x_i) - f(x_{i-1})$

By Lagrange's theorem on finite increments (see Sec. 57), we have

$$\frac{\Delta y_i}{\Delta x_i} = \frac{f(x_i) - f(x_{i-1})}{x_i - x_{i-1}} = f'(\xi_i)$$

where $\xi_i \in (x_{i-1}, x_i)$. Therefore,

$$L_n = \sum_{i=1}^n \sqrt{1 + f'^2(\xi_i)} \Delta x_i$$

The resultant expression is an integral sum for the function $\sqrt{1 + f'^2(x)}$ on the interval $[a, b]$.¹ Passing to the limit as $\max \Delta x_i \rightarrow 0$ (then, of course, the length of the greatest segment of the broken line also tends to zero), we express the length L of the curve by the integral

$$L = \int_a^b \sqrt{1 + f'^2(x)} dx$$

or, briefly,²

$$L = \int_a^b \sqrt{1 + y'^2} dx \quad (*)$$

¹ Since the expression of the function for which this integral sum is formed includes the derivative $f'(x)$ the latter should be *continuous*, which was assumed at the beginning of the present section.

² It should be noted that the element of integration in (*) is the differential of arc length (see Sec. 66). Another method for deriving formula (*) will be given in Sec. 104.

Since $dy = y' dx$, the latter formula can be conveniently rewritten as

$$L = \int_{(A)}^{(B)} \sqrt{dx^2 + dy^2}$$

where (A) and (B) conditionally denote the initial and the terminal points of the arc AB . When computing arc length we substitute for (A) and (B) the corresponding values of the chosen variable of integration.

Now suppose that a smooth curve is represented parametrically by equations $x = x(t)$, $y = y(t)$, and t_1, t_2 ($t_1 < t_2$) are the values of the parameter t corresponding to the end points of the arc. If, as t varies from t_1 to t_2 , the variable point traverses the arc once, moving all the time in one direction, the arc length is given by the formula

$$L = \int_{t_1}^{t_2} \sqrt{x'^2(t) + y'^2(t)} dt \quad (**)$$

If the relations $x = x(t)$, $y = y(t)$ are interpreted as equations of motion of a point and the condition that the point moves in one direction is not introduced formula $(**)$ expresses the path length travelled by the point along the trajectory during time interval from t_1 to t_2 .

Now let a curve be specified by an equation in polar coordinates: $r = r(\varphi)$. Considering the polar angle φ entering into the relations

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

as a parameter we obtain

$$dx = (r' \cos \varphi - r \sin \varphi) d\varphi, \quad dy = (r' \sin \varphi + r \cos \varphi) d\varphi$$

whence

$$\sqrt{dx^2 + dy^2} = \sqrt{r^2 + r'^2} d\varphi$$

Consequently,

$$L = \int_{\alpha}^{\beta} \sqrt{r^2 + r'^2} d\varphi, \quad \alpha < \beta \quad (***)$$

where α and β are the values of the polar angle corresponding, respectively, to the initial point A and the terminal point B of the arc.

The computation of the arc length of a curve is termed the *rectification* of that curve, and a curve possessing arc length is said to be *rectifiable*.

Examples. (1) Let us take a cycloid (Sec. 46)

$$x = a(t - \sin t), \quad y = a(1 - \cos t)$$

and find the length of its arc corresponding to the variation of t from 0 to 2π . The computation gives

$$L = a \int_0^{2\pi} \sqrt{(1 - \cos t)^2 + \sin^2 t} dt = 2a \int_0^{2\pi} \sin \frac{t}{2} dt = 8a$$

and hence the sought-for arc length is eight times the radius of the generating circle.

(2) Let us compute the perimeter of the ellipse

$$x = a \cos t, \quad y = b \sin t \quad (a > b)$$

The quarter ellipse in the first quadrant corresponds to the variation of the parameter t from 0 to $\frac{\pi}{2}$; therefore

$$L = 4 \int_0^{\frac{\pi}{2}} \sqrt{a^2 \sin^2 t + b^2 \cos^2 t} dt = 4a \int_0^{\frac{\pi}{2}} \sqrt{1 - \varepsilon^2 \cos^2 t} dt$$

where $\varepsilon = \frac{\sqrt{a^2 - b^2}}{a}$ is the eccentricity of the ellipse. The substitution $t = \frac{\pi}{2} - \varphi$ leads to the formula

$$L = 4a \int_0^{\frac{\pi}{2}} \sqrt{1 - \varepsilon^2 \sin^2 \varphi} d\varphi$$

The integral on the right-hand side is called the *complete elliptic integral of the second kind*¹; it cannot be computed with the aid of ordinary methods because the antiderivative of the integrand cannot be expressed in terms of elementary functions. Therefore the length of the ellipse is found with the help of tables of values of elliptic integrals. Such tables can be found in many reference books; for example, see [6] and [14].

(3) Let us compute the arc length of the logarithmic spiral $r = ae^{m\varphi}$ from a fixed point $M_0(r_0, \varphi_0)$ to the variable point

¹ The complete elliptic integral of the first kind is the integral
$$\int_0^{\frac{\pi}{2}} \frac{d\varphi}{\sqrt{1 - \varepsilon^2 \sin^2 \varphi}}.$$
 Elliptic integrals are used both in mathematics and its applications for solving many important problems.

$M(r, \varphi)$ (see Sec. 49, Fig. 60). We have

$$\begin{aligned} L &= \int_{\varphi_0}^{\varphi} \sqrt{a^2 e^{2m\varphi} + a^2 m^2 e^{2m\varphi}} d\varphi = a \sqrt{1+m^2} \int_{\varphi_0}^{\varphi} e^{m\varphi} d\varphi = \\ &= a \frac{\sqrt{1+m^2}}{m} (e^{m\varphi} - e^{m\varphi_0}) \end{aligned}$$

Here L is expressed as a function of the polar angle φ of the terminal point. It can also be represented as a function of the polar radius r of that point:

$$L = \frac{\sqrt{1+m^2}}{m} (r - r_0)$$

This formula shows that the arc length of the logarithmic spiral is proportional to the increment of the polar radius of the arc.

If we move along the spiral toward the pole the polar angle φ tends to $-\infty$. The integral expressing the arc length then becomes improper, and we obtain

$$L = a \sqrt{1+m^2} \int_{-\infty}^{\varphi} e^{m\varphi} d\varphi = a \frac{\sqrt{1+m^2}}{m} e^{m\varphi} = \frac{\sqrt{1+m^2}}{m} r$$

i.e. the arc length of the logarithmic spiral from the pole to its arbitrary point is proportional to the polar radius of that point.

The arc length of a smooth space curve specified by equations $x=x(t)$, $y=y(t)$, $z=z(t)$ is computed by the formula analogous to (**):

$$L = \int_{t_1}^{t_2} \sqrt{x'^2(t) + y'^2(t) + z'^2(t)} dt$$

Here t_1 and t_2 are the values of the parameter t corresponding to the end points of the arc, and it is assumed that $t_1 < t_2$ and that as t varies from t_1 to t_2 the variable point describes the whole arc once moving constantly in one direction (see the note after formula (**)).

103. The Centre of Gravity of a Curvilinear Trapezoid. In analytic geometry the division of a line segment in the given ratio is usually demonstrated by the problem of finding the coordinates of the centre of gravity of a system of material points. If masses $m_1, m_2, \dots, m_n$ are concentrated at points $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_n(x_n, y_n)$ in the plane, the coordinates of the centre of gravity $M(\xi, \eta)$ of this system of material points are

en by the well-known formulas

$$\xi = \frac{\sum_{i=1}^n x_i m_i}{\sum_{i=1}^n m_i}, \quad \eta = \frac{\sum_{i=1}^n y_i m_i}{\sum_{i=1}^n m_i}$$

Note that for a homogeneous system of material points, i.e. in case $m_1 = m_2 = \dots = m_n$, the coordinates of the centre of gravity

$$\xi = \frac{\sum_{i=1}^n x_i}{n}, \quad \eta = \frac{\sum_{i=1}^n y_i}{n}$$

independent of the magnitude of the mass and are solely specified by the location of the points.

The product of the mass of a point by its distance from an axis is called the *static moment* of that point about the given axis. For points lying on one side of the axis the distances are taken with plus sign and on the other side with minus sign. Hence, the products $x_i m_i$ and $y_i m_i$ are, respectively, the static moments of the point M_i about Oy and Ox . The sum of the static moments of all the points $M_1, M_2, \dots, M_n$ is called the *static moment of that system of points*. Consequently, the coordinates of the centre of gravity of the system of points $M_i (x_i, y_i) (i = 1, 2, \dots, n)$ can be expressed as

$$\xi = \frac{M_y}{M}, \quad \eta = \frac{M_x}{M} \left(M_x = \sum_{i=1}^n y_i m_i, \quad M_y = \sum_{i=1}^n x_i m_i \right)$$

where M_x and M_y are the static moments of the system about Ox and Oy , and M is its total mass (i.e. the sum of the masses m_i). It follows that

$$\xi M = M_y, \quad \eta M = M_x.$$

Thus, the centre of gravity can be defined as a point such that the total mass of the system is concentrated at it the static moment of that mass point about any axis is equal to the corresponding static moment of the original system.

Now we shall consider a *homogeneous plate (lamina)* of constant thickness. Let the surface density, i.e. the mass per unit area of the plate, be δ kg/m². The centre of gravity of such a plate lies in its middle plane which we take as the plane Oxy . It is obvious that under the given conditions the location of the centre of gravity depends solely on the *geometric form* of the plate. Assume that the region occupied by the plate in the xy -plane is of the

shape of the curvilinear trapezoid,¹ bounded by an arc AB of a curve $y=f(x)$, with base $[a, b]$ on Ox . Let us break up the figure into n subtrapezoids by means of a set of straight lines parallel to the axis of ordinates, the abscissas of the points of intersection of these lines with Ox being $x_0=a, x_1, x_2, \dots, x_n=b$.

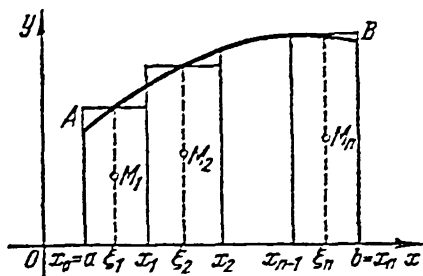


Fig. 126

Replacing the i th subtrapezoid ($i=1, 2, \dots, n$) by the rectangle with altitude equal to the value of the function $f(x)$ at the midpoint ξ_i of the subinterval

$[x_{i-1}, x_i]$ ($\xi_i = \frac{x_{i-1} + x_i}{2}$; see Fig.

126) we obtain a step figure approximating the whole trapezoid.

The location of the centre of gravity of each rectangle is known: it lies at its centre, that

is at the point $M_i \left(\xi_i, \frac{1}{2} f(\xi_i) \right)$, the mass of the rectangle being equal to its area times the density: $\delta f(\xi_i) \Delta x_i$. Now let the mass of each rectangle be concentrated at its centre of gravity. Then the centre of gravity of the resultant system of material points has the coordinates

$$\xi^{(n)} = \frac{\sum_{i=1}^n \xi_i \delta f(\xi_i) \Delta x_i}{\sum_{i=1}^n \delta f(\xi_i) \Delta x_i}, \quad \eta^{(n)} = \frac{\sum_{i=1}^n \frac{1}{2} f(\xi_i) \delta f(\xi_i) \Delta x_i}{\sum_{i=1}^n \delta f(\xi_i) \Delta x_i}$$

Since the density δ is a constant magnitude we can take it outside the summation sign. Then, on cancelling by δ , we receive

$$\xi^{(n)} = \frac{\sum_{i=1}^n \xi_i f(\xi_i) \Delta x_i}{\sum_{i=1}^n f(\xi_i) \Delta x_i}, \quad \eta^{(n)} = \frac{\frac{1}{2} \sum_{i=1}^n f^2(\xi_i) \Delta x_i}{\sum_{i=1}^n f(\xi_i) \Delta x_i}$$

It is clear, geometrically, that if, as usual, n is made to increase indefinitely as all Δx_i 's tend to zero the point with coordinates $(\xi^{(n)}, \eta^{(n)})$ approaches the centre of gravity of the plate. In the denominators of the expressions for $\xi^{(n)}$ and $\eta^{(n)}$ there is an integral sum for the function $y=f(x)$ whose limit is the integral $\int_a^b f(x) dx$, that is, the

¹ The determination of the centre of gravity of a plate of arbitrary shape with variable areal density will be discussed in Chapter VIII, Sec. 132, II.

area S of the plate. In the numerators there are, respectively, integral sums for the functions $xf(x)$ and $f^2(x)$. Consequently,

$$\xi = \frac{\int_a^b xf(x) dx}{S}, \quad \eta = \frac{\frac{1}{2} \int_a^b f^2(x) dx}{S}$$

This can also be written in the abbreviated form

$$\xi = \frac{\int_a^b xy dx}{S}, \quad \eta = \frac{\frac{1}{2} \int_a^b y^2 dx}{S}$$

where $y = f(x)$.

Example. Let us find the centre of gravity of a homogeneous semicircle of radius R . In the coordinate system chosen as shown in Fig. 127 the equation of the upper boundary of the semicircle is written as $y = \sqrt{R^2 - x^2}$. The semicircle being symmetric about Oy , the abscissa ξ of its centre of gravity is equal to zero: $\xi = 0$. Let us determine the ordinate of the centre of gravity:

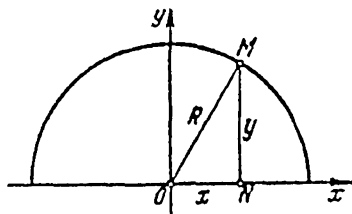


Fig. 127

$$\eta = \frac{\frac{1}{2} \int_{-R}^R (R^2 - x^2) dx}{\frac{1}{2} \pi R^2} = \frac{\left(R^2 x - \frac{x^3}{3} \right) \Big|_{-R}^R}{\pi R^2} = \frac{4}{3\pi} R$$

The method applied here to determining the centre of gravity of a curvilinear trapezoid reduces to dividing it into parts, finding *approximate* location of the centre of gravity of each part and obtaining the coordinates of the centre of gravity of the given trapezoid by passing to the limit in the expressions for the coordinates of the centre of gravity of the corresponding *system of mass points*. This approach is also applicable to finding the centres of gravity of many other physical bodies. As an instance, we shall find the centre of gravity of the solid of revolution generated by the rotation of a given curvilinear trapezoid about the axis of abscissas. Let the volume density of the body be δ kg/m³. Using the notation of Fig. 126 we readily obtain

$$\xi = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n \xi_i \delta \pi f^2(\xi_i) \Delta x_i}{\sum_{i=1}^n \delta \pi f^2(\xi_i) \Delta x_i} = \frac{\int_a^b xy^2 dx}{\int_a^b y^2 dx}$$

(let the reader carry out the calculations). It is apparent that the centre of gravity is on the axis of revolution ($\eta=0$) and therefore is completely specified by the magnitude ξ .

Example. Let us determine the coordinates of the centre of gravity of the hemisphere generated by the rotation of the quarter-circle $y=\sqrt{R^2-x^2}$ ($0 \leq x \leq R$) about the axis of abscissas:

$$\xi = \frac{\int_0^R x(R^2-x^2) dx}{\int_0^R (R^2-x^2) dx} = \frac{\left(R^2 \frac{x^2}{2} - \frac{x^4}{4}\right)\Big|_0^R}{\left(R^2x - \frac{x^3}{3}\right)\Big|_0^R} = \frac{\frac{R^4}{4}}{\frac{2R^3}{3}} = \frac{3}{8}R$$

(the coordinate η is of course equal to zero).

Let the reader derive the following formulas for the coordinates of the centre of gravity of the *homogeneous line* specified by an equation $y=f(x)$, the initial point A and the terminal point B of the line having, respectively, the abscissas equal to a and b :

$$\xi = \frac{\int_a^b x \sqrt{1+y'^2} dx}{L} = \frac{\int_a^b x ds}{L}, \quad \eta = \frac{\int_a^b y \sqrt{1+y'^2} dx}{L} = \frac{\int_a^b y ds}{L}$$

where L is the length of the arc AB and ds is the differential of arc length.

§ 2. General Scheme of the Application of the Integral

104. The Problem-solving Scheme. In the problems we have discussed the general scheme of the application of the definite integral is essentially the same. The desired magnitude (such as area, work, path length, volume, etc.) considered in the problem corresponds to an interval of variation of a variable which serves as a variable of integration: the area of a trapezoid corresponds to an interval $[a, b]$ of variation of the abscissa x , the work corresponds to an interval $[s_0, s]$ of variation of path length travelled, the path s to an interval $[T_0, T]$ of variation of time t , etc.

Let us describe this scheme in the general form. We shall suppose that the variable we spoke about is denoted by x and that its interval of variation is $[a, b]$.

The first step is to break up the interval $[a, b]$ into subintervals $[x_{i-1}, x_i]$; $i=1, 2, \dots, n$ ($x_0=a, x_n=b$). This is done because of the essential assumption that *the value of the magnitude in question corresponding to the whole interval is representable as the sum of its values corresponding to the subintervals* (the so-called *additivity property*). Indeed, area, work, path length, etc.

possess additivity with respect to the intervals they are related to: the area of a curvilinear trapezoid with an interval as base is equal to the sum of the areas of the subtrapezoids whose bases constitute the whole interval, the work performed by a force on a given path is equal to the sum of the works on the subintervals the whole path is divided into and so on.

The next step is to form the corresponding sum (which is an integral sum) giving an approximation to the sought-for value of the magnitude in question; the greater the number n of the subintervals and the smaller the maximum subinterval, the more precise the approximation.

Finally, on passing to the limit as $n \rightarrow \infty$ we find the required value in the form of the integral

$$I = \int_a^b f(x) dx \quad (*)$$

where $f(x)$ is the function specified by the conditions of the problem (for instance, the ordinate of the line bounding the given trapezoid, the acting force and the like).

This scheme is also applicable to any other problem of that type. But now we are going to show that the approach considered above (i.e. breaking up the interval into parts, summation, and passage to the limit) is equivalent to a simpler standard technique proceeding from the form of the final result (the sought-for magnitude is equal to the integral of a certain function) and based on the theory of the definite integral.

Let us consider the value of the magnitude we are interested in (i.e. area, work, path length, etc.) corresponding to a *variable* interval $[\alpha, x]$ whose left end point is fixed: $\alpha = \text{const}$ ($\alpha \leq a$). Then this value becomes an unknown function $u = F(x)$ of the abscissa of the right end point x of the variable interval. If the function $F(x)$ were known it would be easy to determine the desired value of the magnitude corresponding to the *given interval* $[a, b]$ (denote this value by I). For, $F(b)$ corresponds to the interval $[\alpha, b]$ and $F(a)$ to $[\alpha, a]$ while I to the interval $[a, b]$, and hence, by the *additivity* of the integral, we have

$$F(b) = F(a) + I$$

that is

$$I = F(b) - F(a) \quad (**)$$

Relation (**) becomes particularly visual if we take the example of the area of the curvilinear trapezoid (Fig. 128). Here we obviously have

$$I = S_{aABb}, \quad F(a) = S_{\alpha CAa} \quad \text{and} \quad F(b) = S_{\alpha CBb}$$

The function $F(x)$ is supposed to be differentiable, and, by Newton-Leibniz formula,

$$I = F(b) - F(a) = \int_a^b dF(x)$$

that is the sought-for value I of the magnitude is equal to the integral of the differential of the function $u = F(x)$ taken over the interval $[a, b]$ to which it corresponds.

Thus, the element of integration $f(x)dx$ in formula (*) is nothing but the differential du of the function $u = F(x)$, and, consequently,

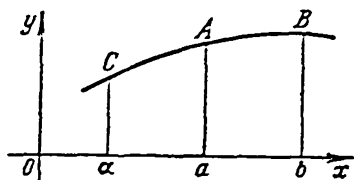


Fig. 128

in order to solve the problem it only suffices to know the differential of the function $F(x)$. After the differential has been found the sought-for value of the magnitude is expressed by the integral of that differential taken over the interval $[a, b]$.

Now, by analogy with the concrete problems we have dealt with,

it is possible to give the general description of the simplest model of the problem solvable with the aid of the definite integral. The basic features of this characteristic problem are:

(1) It is required to determine the value I of a magnitude u corresponding to a given interval $[a, b]$ of variation of an independent variable x . The magnitude u possesses additivity, that is, if the interval is divided into parts its value corresponding to the entire interval is equal to the sum of the values corresponding to the subintervals.

(2) If the magnitude u is related to a variable interval whose end point α (for definiteness, the left one) is fixed ($\alpha = \text{const}$, $\alpha \leq a$) while the other end point is variable and has the abscissa x , it is possible to consider u as a function of the variable x , i.e. $u = F(x)$. The function $F(x)$ is supposed to be differentiable.

If both conditions hold the value I of the magnitude is expressed by the integral

$$I = F(b) - F(a) = \int_a^b dF(x)$$

and thus, the problem completely reduces to finding the differential $du = dF(x)$.

Although in the general case the function $u = F(x)$ is unknown its differential can often be determined on the basis of the conditions of the problem. To this end, we take an arbitrary value of x belonging to the interval $[a, b]$ and give it an infinitesimal increment dx . The part of the corresponding increment $\Delta F(x)$ of

the function which would appear if all the other parameters specifying the expression of $F(x)$ retained in the interval $[x, x+dx]$ their values assumed at the point x is usually the differential $dF(x)$. In every concrete situation we can always check whether the expression thus found is the differential by verifying that it is proportional to dx and that it differs from $\Delta F(x)$ by an infinitesimal of higher order than dx .

The last requirement is equivalent to the condition that the limit of the ratio of the expression $dF(x)$ thus obtained to $\Delta F(x)$ tends to unity as $dx \rightarrow 0$. What has been said is demonstrated by the example below.

Let us take the simplest problem of finding the area of a curvilinear trapezoid whose solution is already known and obtain this solution with the aid of the new technique described above. Let the trapezoid be bounded above by a line $y=f(x)$; its area corresponding to the interval $[a, b]$ of variation of x will be denoted by S . The area of the trapezoid αCMx (see Fig. 129) is

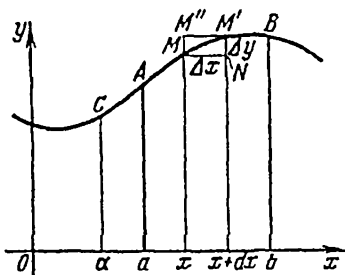


Fig. 129

a function of x which we denote by $S(x)$. Let us choose an arbitrary value of x and give it an increment dx . Then the increment $\Delta S(x)$ is represented by the area of the trapezoid $xMM'(x+dx)$. Now we shall prove that the area of the rectangle $xMM'N(x+dx)$, that is, the part of the increment $\Delta S(x)$ appearing if the function $f(x)$ is regarded as constant in the interval $[x, x+dx]$, is the differential of the function $S(x)$:

$$dS(x) = f(x) dx$$

The expression of $dS(x)$ given by this formula is proportional to dx , and it only remains to check that $\Delta S(x)$ and $f(x)dx$ are equivalent infinitesimals. As is seen in Fig. 129, their difference does not exceed the area of the rectangle $MM'M'N$ which is equal to the product $\Delta x \Delta y$:

$$\Delta S(x) - dS(x) < \Delta x \Delta y \quad (dS(x) = f(x) dx)$$

This difference is thus an infinitesimal of higher order than $\Delta x = dx$, which completes the proof of our assertion (see Sec. 51, IV). By what has been proved, we have

$$S = \int_a^b f(x) dx$$

Consider another example. Since we know (see Sec. 66 of Chapter IV) that the differential of arc length ds is given by the for-

mula

$$ds = \sqrt{1 + y'^2} dx$$

we can immediately write (cf. Sec. 102) the expression for the arc length L :

$$L = \int_a^b \sqrt{1 + y'^2} dx$$

105*. The Area of a Surface of Revolution. In this section we shall derive a formula for computing the area of a surface of revolution.

Let an arc AB of a curve $y = f(x)$ be rotated about Ox (Fig. 130). It is required to find the area Q of the surface of revolution thus generated on condition that the function $f(x)$ possesses a continuous derivative.

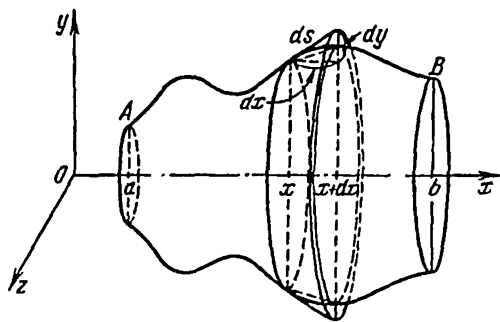


Fig. 130

Let $Q(x)$ designate the area corresponding to the interval $[a, x]$; our task is to find its differential. To this end we make the following construction. Draw through an arbitrary point x ($a < x < b$) a plane perpendicular to Ox . Give x an increment dx and draw

another plane perpendicular to Ox and passing through the point $x + dx$. The area $Q(x)$ of the surface then gains an increment ΔQ equal to the area of the portion of the surface contained between these parallel planes passing through the points x and $x + dx$ perpendicularly to Ox . The magnitude ΔQ satisfies the following obvious inequalities (in accordance with Fig. 130 we assume that $\Delta y < dy$):

$$\frac{2\pi y + 2\pi(y + \Delta y)}{2} \sqrt{dx^2 + \Delta y^2} < \Delta Q < \frac{2\pi y + 2\pi(y + dy)}{2} \sqrt{dx^2 + dy^2}$$

The leftmost expression is equal to the lateral area of the frustum of the cone whose slant height is the chord of the arc while the rightmost expression is the area of another cone with the segment of the tangent line as slant height.¹ On dividing all the members of the inequalities by $2\pi y \sqrt{dx^2 + dy^2}$ we obtain

$$\frac{4\pi y + 2\pi \Delta y \sqrt{dx^2 + \Delta y^2}}{4\pi y \sqrt{dx^2 + dy^2}} < \frac{\Delta Q}{2\pi y \sqrt{dx^2 + dy^2}} < \frac{4\pi y + 2\pi dy}{4\pi y}$$

¹ Here we suppose that on the interval $[x, x + dx]$ the arc of the curve is convex (upward). These inequalities are evident from the geometrical point of view but its analytical aspects are not justified rigorously. The detailed theoretical justification of the formula for the surface area of a solid of revolution involves some more sophisticated considerations.

Let the reader show that both leftmost and rightmost members of the latter relations tend to 1 as $dx \rightarrow 0$; it follows that

$$\frac{\Delta Q}{2\pi y \sqrt{dx^2 + dy^2}} \rightarrow 1$$

Consequently, the expression

$$2\pi y \sqrt{dx^2 + dy^2} = 2\pi y \sqrt{1 + y'^2} dx$$

is nothing but the differential $dQ(x)$, whence

$$Q = 2\pi \int_a^b y \sqrt{1 + y'^2} dx$$

or, in the abbreviated form,

$$Q = 2\pi \int_a^b y ds$$

where ds is the differential of arc length.

These formulas make it possible to pass to the case when the line whose revolution generates the surface is specified parametrically or by an equation in polar coordinates. Let the reader derive the corresponding formulas.

Example. Let us compute the area of a spherical segment generated by the rotation of an arc of a circle, with centre at the origin and radius r , about Ox .

The equation $x^2 + y^2 = r^2$ of the circle implies $y^2 = r^2 - x^2$ and $yy' = -x$; hence,

$$Q = 2\pi \int_a^b \sqrt{r^2 - x^2} \sqrt{1 + \frac{x^2}{r^2 - x^2}} dx = 2\pi \int_a^b r dx = 2\pi r (b - a) = 2\pi r H$$

where H is the altitude of the spherical segment. For $H = 2r$ we obtain the surface area of the sphere: $Q = 4\pi r^2$.

106. Fluid Pressure on the Wall of the Vessel. Let a fluid with specific weight γ N/m³ fill a vessel of the form of a rectangular parallelepiped. We are going to determine the force of pressure P of the fluid acting upon one of the walls of the vessel, the base of that wall being a and altitude h . This problem admits of the application of the general scheme elaborated in Sec. 104 since the force of pressure on the whole wall is the sum of the pressures on the parts it is divided into.

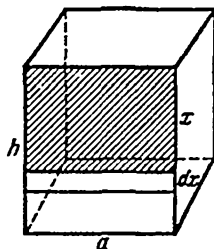


Fig. 131

Let $P(x)$ designate the pressure on the shaded part of the wall shown in Fig. 131. We now give x an increment dx and compute the principal part, proportional to dx , of the corresponding incre-

ment $\Delta P(x)$ assuming that the depth of the layer of the fluid between x and $x+dx$ is (approximately) constant and equal to x . Then

$$dP(x) = \gamma x dx \quad (*)$$

Since the increment of the force of pressure satisfies the inequality $\Delta P(x) < \gamma a(x+dx)dx$, the difference between $\Delta P(x)$ and $dP(x)$ does not exceed

$$\gamma a(x+dx)dx - \gamma x dx = \gamma a dx^2$$

that is, it is an infinitesimal of higher order than dx .

Consequently, we have shown that expression (*) is in fact the differential of the function $P(x)$. The solution of the problem is now completed by the integration from 0 to h :

$$P = \int_0^h \gamma a x dx = \frac{\gamma a h^2}{2}$$

It is interesting to note that the formula we have derived shows that the pressure on the wall is equal to the force of pressure acting upon a horizontal plane area of the same dimensions immersed in the fluid at the depth of the centre of gravity of the wall.

QUESTIONS

1. How is the area of a plane figure found in Cartesian coordinates, polar coordinates and in the case when its boundary is represented by parametric equations?

2. Derive the formula for computing the volume of a body when the areas of its parallel sections are known.

3. Derive the formula for the volume of a solid of revolution.

4. State the definition of the arc length of a curve. Write down the formulas for computing arc length in Cartesian coordinates, polar coordinates and for the case of parametric equations of the curve.

5. Derive the formulas for the coordinates of the centre of gravity of a curvilinear trapezoid.

6. Describe the general scheme for application of the definite integral to solving geometrical and physical problems. Characterize the basic features of the problems solvable with the aid of the integral.

7. Give examples of the application of the scheme.

8*. Write down the formulas for computing the area of the surface of revolution generated by the rotation of a curve specified by an equation in Cartesian coordinates, polar coordinates and also for parametric equations of the generating curve.

Chapter VII

FUNCTIONS OF SEVERAL VARIABLES AND THEIR DIFFERENTIATION

§ 1. Functions of Several Variables

107. Functions of Two Variables. Method of Parallel Sections. Domains and Neighbourhoods. In the foregoing chapters we studied functions of one independent variable. But there are many cases when a magnitude depends not on one but on two or more independent variables, that is, when the values of the variable magnitude in question are specified by the values of several independent magnitudes. In such cases we speak about a function of two or more arguments.

For instance, the area S of a rectangle is a function of two independently varying quantities which are its sides a and b ; this function is expressed by the formula

$$S = ab$$

Similarly, the volume V of a rectangular parallelepiped is a function of three independent variables which are the lengths a , b and c of the edges of the parallelepiped:

$$V = abc$$

Another example is the work A of a direct electric current flowing through a section of a circuit; it depends on the voltage U across the ends of the section, on the intensity I of the current and on time t . This functional relationship is given by the formula

$$A = IUt$$

1. Functions of two variables. We shall begin with the case of two independent variables which will be denoted x and y .

To every pair of values of x and y there corresponds a point in the xy -plane for which they serve as coordinates. Let us take a set of points D in the plane Oxy .

Definition. A variable z is said to be a *function of two variables x and y defined on the set D* if to each point of the set there corresponds a definite value of the variable z .

The point set D is then referred to as the *domain of definition of the function*. The domain of definition of a function is usually a part of the xy -plane bounded by one or several lines.

If a magnitude z is a function of variables x and y we write

$$z = f(x, y)$$

In this relation the letter f symbolizing the function (it can be replaced by any other symbol) is followed by the letters designating the independent variables (the *arguments*) which are placed into the parentheses and separated by the comma.

A function of two variables, like a function of one variable, can be specified by a *table*, by a *formula* (i. e. *analytically*) or by its *graph* (see Sec. 8).

A *tabular representation* of a function indicates the values of the function for a number of pairs of values of the independent variables. For this purpose we usually compile a table of double entry; this can be illustrated by the table below specifying the dependence of the screw gear efficiency η on the coefficient of friction μ and on the helix pitch angle α :

$\mu \backslash \alpha$	$\frac{\pi}{36}$	$\frac{\pi}{18}$	$\frac{\pi}{12}$
0.01	0.897	0.945	0.961
0.02	0.812	0.895	0.925
0.03	0.743	0.850	0.892

In the *analytical specification* of a function we use a formula determining the values of the function depending on the values of the independent variables.

For example, each of the formulas

$$z = 2x + 3y - 5, \quad z = \frac{xy}{x^2 + y^2} \quad \text{and} \quad z = \frac{\sin(2x + 3y)}{\sqrt{1 + (x + y)^2}}$$

specifies z as a function of x and y . As a rule, when a function is represented analytically its domain of definition is understood (unless there are some additional conditions) as the maximum range of the variables x and y for which the formula specifying the function makes sense, that is, for the values of x and y which can be substituted into the formula to obtain the corresponding real values of z . For instance, the domain of the function

$$z = \sqrt{r^2 - x^2 - y^2}$$

is the set of points in the xy -plane whose coordinates satisfy the relation

$$x^2 + y^2 \leq r^2$$

that is, the *circle* of radius r with centre at the origin. Similarly, for the function

$$z = \ln(x^2 + y^2 - r^2)$$

the domain of definition is the set of points whose coordinates satisfy the condition

$$x^2 + y^2 > r^2$$

that is, the *exterior* of the circle.

The function $z = \frac{x}{1+x^2+y^2}$ is defined throughout the plane Oxy while the function $z = \frac{\sin(xy)}{x-y}$ is defined in the plane with the straight line $y = x$ deleted.

As in the case of a function of one argument, we can consider *implicit* functions of two independent variables. An equation of the form

$$F(x, y, z) = 0$$

usually specifies each of the variables x, y, z involved as an implicit function of the other two variables; this question will be discussed in more detail in Sec. 117.

If z is a function of two independent variables x and y its *graph* (in Cartesian coordinates) is the set of points whose abscissas and ordinates are the values of x and y and the third coordinate is the corresponding value of z . The graph of a function defined in a region of the plane is usually a *surface*.

As an example, consider the equation $x^2 + y^2 + z^2 = R^2$. It specifies the sphere of radius R with centre at the origin, and therefore the graphs of the functions

$$z = \sqrt{R^2 - x^2 - y^2} \text{ and } z = -\sqrt{R^2 - x^2 - y^2}$$

are, respectively, the upper and the lower hemispheres (Fig. 132).

The graph of the function

$$z = x^2 + y^2$$

is a paraboloid of revolution (Fig. 133) and that of the linear function

$$z = ax + by + c$$

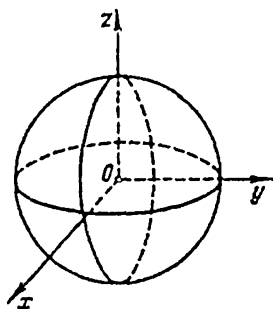


Fig. 132

is a plane; in particular, the graph of a constant, that is, of the function $z=C$ ($C=\text{const}$) is a plane parallel to Oxy . Conversely, a surface in the xyz -space represents a function of two independent variables whose values are equal to the z -coordinates of the points of that surface corresponding to the values of the independent variables equal to the abscissas and the ordinates.

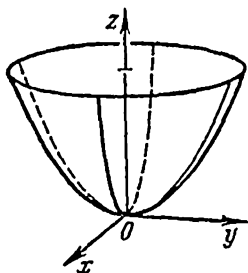


Fig. 133

To represent a function *graphically* means to specify its graph.

II. The method of parallel sections. Level lines. In analytic geometry, when studying second-order algebraic surfaces we apply the *method of parallel sections* in which the shape of the surface is analysed with the aid of its equation by investigating the curves formed in the intersections of the surface with planes parallel to the coordinate planes. This

method is also applicable to studying arbitrary functions of two variables. For instance, suppose we are given a function

$$z = f(x, y)$$

specifying a surface in the xyz -space. If a value y_0 of the argument y is fixed while x remains variable the coordinate z becomes a function of one independent variable x :

$$z = f(x, y_0)$$

The character of variation of the variable z as function of x can now be investigated by applying to the above function the known methods for studying functions of one argument.

This means, geometrically, that we consider the line of intersection of the surface $z = f(x, y)$ with the plane $y = y_0$ parallel to the xz -plane (see the line CD in Fig. 134). Giving y another constant value y_1 we obtain the line C_1D_1 , etc.

The behaviour of z as y is varied and x is given different constant values can be investigated in a similar way. For instance, in Fig. 134 we see

the line of intersection of the surface $z = f(x, y)$ with the plane $x = x_0$ (the line AB). To investigate this line we should study the function of one variable $z = f(x_0, y)$. If the disposition of these lines is known we can judge upon the shape of the surface.

The investigation of the given function $z = f(x, y)$ by means of functions of one argument can be carried out in another way

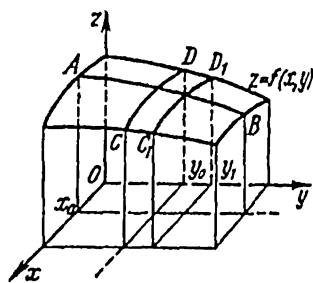


Fig. 134

by fixing the values of the function itself instead of the values of independent variables. To this end, let us put $z = z_0$; this results in the equation

$$f(x, y) = z_0 \quad (*)$$

determining a relationship between the variables x and y , that is, a function of one argument, corresponding to the given fixed value z_0 of the function z . Geometrically, when a value z_0 of z is fixed we obtain the line of intersection of the surface $z = f(x, y)$ with the plane $z = z_0$ parallel to the plane Oxy .

Equation (*) specifies the projection l of the line of intersection L of the surface $z = f(x, y)$ with the plane $z = z_0$ (see Fig. 135).¹ When the point with coordinates x and y moves along the line l the function retains the constant value z_0 .

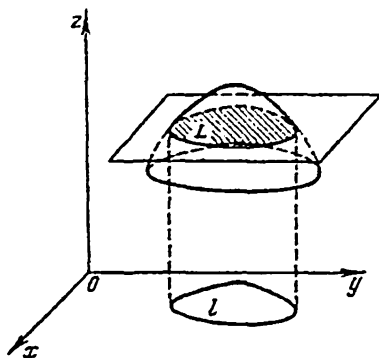


Fig. 135

Definition. A curve in the xy -plane at whose all points of the function $z = f(x, y)$ retains a constant value is called a *level line* of the function.

Taking different values $z = z_0$ we obtain a family of level lines of the function $z = f(x, y)$; if the difference between neighbouring values of z_0 is sufficiently small such a family enables us to visualize the behaviour of the function.

Let us make z assume the values $\dots -3h, -2h, -h, 0, h, 2h, 3h, \dots$ where h is a positive number; we then obtain a system of *evenly spaced* parallel sections and the corresponding family of level lines which, in this case, when the cutting planes are *equidistant apart*, are termed *contour lines*. In the parts of the xy -plane where the contour lines are thicker the function varies "rapidly" and in the parts where they are rarer it varies "slowly". For, in the former case to the increment h of the function there corresponds a smaller displacement of the point $P(x, y)$ in the plane than in the latter case.

In Fig. 136 we see a family of contour lines with $h = 1$ constructed for the function $z = x^2 + y^2$ (representing a paraboloid of revolution). These contour lines are concentric circles with centre at the point $O(0, 0)$ and that point itself. It is clearly seen that, as z increases, the circles become closer to each other, which means that the surface rises more steeply.

¹ The equation $f(x, y) = z_0$ specifies a cylindrical surface in the xyz -space with elements parallel to Oz and directrix l (or L).

In applied sciences contour lines are frequently used for representing the behaviour of functions of two variables. For example, in cartography the altitude of a terrain above sea level at a given point considered as a function of two variables (the coordinates of that point) is represented by contour lines of this function.

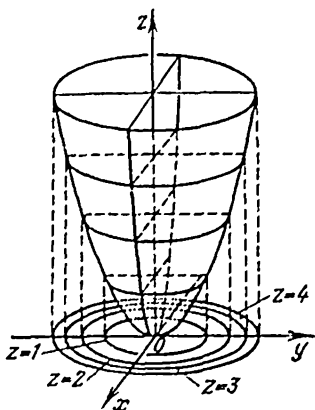


Fig. 136

These lines called *horizontals* are extremely convenient for representing the altitude of the terrain. In meteorology the so-called *isotherms* and *isobars* are used; they are, respectively, lines of constant average temperature and of constant average pressure and represent these magnitudes as functions of the coordinates of the point of the terrain.

III. The domain of definition of a function of one variable is most often an interval of the number line. For a function of two variables the domain of definition is a set of points in the coordinate plane. Let us characterize the point sets of this kind which will be most frequently considered in what follows. We shall use visual geometrical interpretation; to begin with, let us state the definition of an *r*-neighbourhood of a point in the plane.

Definition. An *(open) r*-neighbourhood of a given point $P_0(x_0, y_0)$ is the set of all points lying inside the circle of radius *r* and centre at the point P_0 (Fig. 137).

The coordinates of these points satisfy the inequality

$$(x - x_0)^2 + (y - y_0)^2 < r^2$$

If the points of the circumference of the circle are added to the *r*-neighbourhood we obtain the *closed r*-neighbourhood whose points have coordinates satisfying the non-strict inequality $(x - x_0)^2 + (y - y_0)^2 \leq r^2$.

The basic geometrical figure in the plane is a *domain*:

An *open domain in the plane* is a point set *D* satisfying the following two conditions:

- (1) If a point belongs to the set *D* there exists a neighbourhood of this point whose all points also belong to *D*.
- (2) Any two points of the set *D* can be joined with a continuous line lying entirely within *D*.

The first condition means that all the points of an open domain are *interior*, i. e. every point of the domain lies inside it together

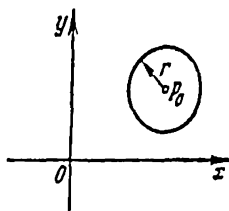


Fig. 137

with an r -neighbourhood of that point. The second condition indicates that an open domain is (*arcwise*) *connected*: its any two points can be connected by a line whose all points are in the domain.

Examples of open domains are the interiors of a circle, of an ellipse, of a polygon, of an annulus, etc. (see Fig. 138). In what follows we shall always deal with domains bounded by finite arcs of smooth curves. These arcs form the *boundaries* of the domains.

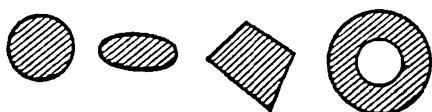


Fig. 138

If all the points of the boundary of an open domain D are added to D we obtain a *closed* domain. When speaking about a domain D we usually mean an open domain; the case of a closed domain (often denoted $\overline{D}$) will always be stipulated.

Note that a region of the shape of a figure "8" (see Fig. 139) is not a domain since the points lying within different loops cannot be joined by a continuous line whose all points are interior.



Fig. 139

A domain is said to be *bounded* (*finite*) if it entirely lies inside a circle with centre at the origin; thus, the distance from any point of a bounded domain to the origin does not exceed a fixed number. All the domains mentioned above are bounded. Examples of *unbounded* domains are a *half-plane* (i. e. a part of the plane lying on one side of a given straight line), a *strip* (a part of the plane enclosed within two parallel straight lines), the *exterior* of any circle and the like. Any two intersecting straight lines break up the plane into four unbounded angular domains ("*wedges*").

In what follows an important role will be played by so-called *simply connected* (1-connected) domains:

A domain is said to be *simply connected* (or 1-connected) if all the points of the plane lying inside a simple closed curve whose all points belong to the domain are also within the domain. A curve is called *simple* if it does not intersect itself (for instance, the boundary of a circle is a simple curve but that of a figure "8" is not).

Examples of simply connected domains are the interiors of a circle and of an ellipse; a half-plane, a strip, a wedge, etc. Any domain bounded by a simple closed curve is simply connected.

As an example of a *not simply connected* (*multiply connected*) domain is an annulus (see Fig. 140a) since not all points of the plane enclosed by a circle Γ lying within the annulus belong to this annulus. The boundary of an annulus consists of two nonintersecting curves; accordingly, we say that an annulus is a *doubly connected* (*2-connected*) domain. The domain shown in Fig. 140b is *triply connected* (*3-connected*); its boundary consists of three nonintersecting curves. In the general case the notion of an *n-tuply connected* domain is introduced in a similar way.

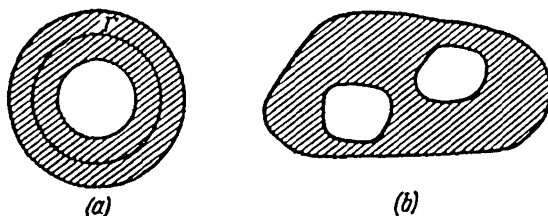


Fig. 140

108. Limit of a Function of Two Variables. Continuity.

1. The limit of a function. The definition of a limit of a function of two variables is stated by analogy with the case of a function of one variable:¹

Definition. A number A is called the *limit* of the function $z = f(x, y)$ as $x \rightarrow x_0$, $y \rightarrow y_0$ if for all the values of x and y which are, respectively, sufficiently close to the numbers x_0 and y_0 , the corresponding values of the function $f(x, y)$ are arbitrarily close to the number A .

If A is the limit of $f(x, y)$ as $x \rightarrow x_0$, $y \rightarrow y_0$ we write

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = A$$

In this definition it is not supposed that the function is defined at the point $P_0(x_0, y_0)$ itself, and therefore it is assumed that at least one of the inequalities $x \neq x_0$, $y \neq y_0$ is fulfilled, that is the point (x, y) does not coincide with the limiting point (x_0, y_0) .

This definition can be re-stated in terms of inequalities: If, given an arbitrary $\epsilon > 0$, there exists a number $\delta > 0$ such that for all the points $P(x, y)$ whose coordinates satisfy the inequality

$$0 < (x - x_0)^2 + (y - y_0)^2 < \delta^2$$

(that is, for all points $P \neq P_0$ belonging to the δ -neighbourhood of

¹ Before studying the subject of this section the reader should recall the content of Sec. 24, Chapter II.

the point P_0) the inequality

$$|f(x, y) - A| < \epsilon$$

is fulfilled, the number A is the limit of the function $f(x, y)$ as $x \rightarrow x_0, y \rightarrow y_0$.

The definition of the limit is readily extended to the case when one of the independent variables or both tend to infinity. For instance, the relation $\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow +\infty}} f(x, y) = A$ means that for any $\epsilon > 0$

there are numbers $\delta > 0$ and $N > 0$ such that the inequalities $0 < |x - x_0| < \delta$ and $y > N$ imply the inequality $|f(x, y) - A| < \epsilon$. It is of course assumed that the function $f(x, y)$ is defined for all the values of x and y in question. Let the reader state the definition of the limit for the other possible cases when the independent variables tend to infinity.

The definitions of infinitesimals and of infinitely large magnitudes which are functions of two variables are analogous to the corresponding definitions for functions of one variable, and we shall not state them here.

It should also be noted that all the rules for passing to the limit given in Sec. 29, Chapter II for functions of one variable are extended without any changes to the case of functions of several arguments.

II. Continuity. Let a function $f(x, y)$ be defined in a domain D and let $P_0(x_0, y_0)$ be an interior point of the domain. The increment of the function $z = f(x, y)$ at the given point P_0 (recall the definition of the increment for a function of one variable) corresponding to given increments Δx and Δy of the independent variables is the difference

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

(it is of course supposed that the point $(x_0 + \Delta x, y_0 + \Delta y)$ also belongs to the domain D).

Definition. A function $z = f(x, y)$ is said to be *continuous at a point* (x_0, y_0) if it is defined in a neighbourhood of that point and if to infinitesimal increments of x and y there corresponds an infinitesimal increment of z , that is,

$$\lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \Delta z = 0$$

Denoting $x_0 + \Delta x$ by x and $y_0 + \Delta y$ by y we can re-write the condition of the continuity of the function $f(x, y)$ in the form

$$\lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} [f(x, y) - f(x_0, y_0)] = 0 \quad \text{or} \quad \lim_{\substack{x \rightarrow x_0 \\ y \rightarrow y_0}} f(x, y) = f(x_0, y_0)$$

Thus, if a function is continuous at a point (x_0, y_0) its limit at

that point coincides with the particular value of the function at this point; the converse is also true.

We say that a function is continuous in a domain D if it is continuous at every point of the domain.

If $P_0(x_0, y_0)$ is a boundary point of the domain D the definition of the continuity of the function at that point should be slightly changed. Namely, in this case the limit of the function $f(x, y)$ as $x \rightarrow x_0, y \rightarrow y_0$ must be understood in the sense similar to the case of one-sided limit of a function of one variable: the point (x, y) , when approaching the point (x_0, y_0) , should move within the domain D , that is, the definition of the limit must involve not all the points (x, y) lying sufficiently close to (x_0, y_0) but only those belonging to the domain D .

From the geometrical point of view the continuity of the function $z = f(x, y)$ means that its graph is a surface without discontinuities.

The basic properties of continuous functions of two independent variables are the same as in the case of functions of one variable (see Sec. 35, Chapter II).

A point belonging to the domain D or to its boundary at which the condition of continuity is violated is called a *point of discontinuity* of the function¹. Points of discontinuity of a function of two variables may form arcs or curves.

Let us consider examples of discontinuous functions.

The function $z = \frac{1}{x^2 + y^2}$ is continuous everywhere except at the point $(0, 0)$ which is its point of discontinuity. It is clear that $z \rightarrow \infty$ as $x \rightarrow 0, y \rightarrow 0$. This means, geometrically, that the graph of the function is a surface having an infinite "spire" at the point $(0, 0)$.

Recalling the general equation of a surface of revolution derived in analytic geometry we conclude that the equation $z = \frac{1}{x^2 + y^2}$ specifies the surface generated by the rotation of the curve $z = \frac{1}{x^2}, y = 0$ lying in the xz -plane about Oz .

The points of discontinuity of the function $z = \frac{1}{x^2 - y^2}$ are all the points of the bisectors of the quadrants of the xy -plane, that is, those points for which $y = x$ or $y = -x$.

A more complicated example is the point of discontinuity $(0, 0)$ for the function

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

¹ A more precise definition of a point of discontinuity of a function of two variables is analogous to the corresponding definition for a function of one variable (see Sec. 33).

At each point of the straight line $y=kx$ where k is an arbitrary number the function assumes the constant value

$$f(x, y) = \frac{kx^2}{x^2 + k^2x^2} = \frac{k}{1+k^2}$$

If the point (x, y) approaches the point $(0, 0)$ along the line $y=kx$ the limit of the values of the function is equal to $\frac{k}{1+k^2}$. Making k take on different values, i. e. making (x, y) move toward $(0, 0)$ along different straight lines, we obtain *different* limits of the values of the function $f(x, y)$. This means that the function has no limit at the point $(0, 0)$ as x and y approach the origin *in an arbitrary way* and thus is discontinuous at that point.

109. Functions of Several Independent Variables. Here we shall generalize the definitions stated in the foregoing sections to the case of more than two independent variables.

If there are three independent variables they are usually denoted by x, y and z . If u is a function of x, y, z we write

$$u = f(x, y, z)$$

With each triple of values x, y, z we can associate the point (x, y, z) of the xyz -space; then the domain of definition of the function $f(x, y, z)$ can be interpreted geometrically as a part of the three-dimensional space (it usually is a spatial region bounded by a surface). The definition of a three-dimensional domain is completely analogous to that of a plane domain; the only change is that *an r -neighbourhood of a point $P_0(x_0, y_0, z_0)$ is in this case understood as the set of all interior points of the ball of radius r with centre at the point P_0 .* The coordinates of the points of this neighbourhood satisfy the inequality

$$(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2 < r_0^2$$

The "graph" of a function of three variables $u=f(x, y, z)$ should be understood (formally) as the collection of the 4-tuples (x, y, z, u) such that x, y and z run through the domain of definition of the function while u assumes the corresponding values of the function $f(x, y, z)$; it cannot of course be visualized geometrically since we live in the three-dimensional geometrical space and cannot draw the fourth coordinate axis Ou perpendicular to given mutually perpendicular axes Ox, Oy and Oz .

If the number n of the independent variables exceeds three it is inconvenient to denote them by different letters, and therefore they are most often denoted by one letter, for instance, x , supplied with different subscripts: $x_1, x_2, \dots, x_n$. If u is a function of these variables we write

$$u = f(x_1, x_2, \dots, x_n)$$

To designate concrete numerical values of the independent variables $x_1, \dots, x_n$ we should use, instead of ordinary subscripts, some other symbols, for instance $x_1^0, x_2^0, \dots, x_n^0$.

For the sake of convenience we shall (formally) retain the ordinary geometrical terminology: an n -tuple $(x_1, x_2, \dots, x_n)$ will be called a point of the n -dimensional space, the latter being understood as the collection of all such n -tuples. We shall not state the general definition of a domain in the n -dimensional space of n -tuples and shall confine ourselves to the following definition of an r -neighbourhood: *the set of points $(x_1, x_2, \dots, x_n)$ of the n -dimensional space whose coordinates satisfy the inequality*

$$(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \dots + (x_n - x_n^0)^2 < r^2$$

is called the r -neighbourhood of the point $P_0(x_1^0, x_2^0, \dots, x_n^0)$. Such a neighbourhood is referred to as an *open n -dimensional ball*, and its boundary, that is, the set of points whose coordinates satisfy the equation

$$(x_1 - x_1^0)^2 + (x_2 - x_2^0)^2 + \dots + (x_n - x_n^0)^2 = r^2$$

is called an *n -dimensional sphere* (or, simply, an *n -sphere*).

The definition of the limit of a function and that of the continuity of a function for the case of n independent variables are completely analogous to the corresponding definitions for functions of two variables, and we shall not state them here.

§ 2. Derivatives and Differentials. Differential Calculus

110. Partial Derivatives and Differentials.

I. Partial derivatives. Let $z = f(x, y)$ be a function of two independent variables x and y . We begin with fixing a constant value of the argument y and investigating the function of one variable x thus appearing. Suppose that this function (we write it in the original form $f(x, y)$ where $y = \text{const}$) possesses a derivative with respect to x ; this derivative is equal to

$$\lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

We shall denote this limit as $f'_x(x, y)$ where the subscript indicates the variable x with respect to which the derivative is taken for a fixed value of y .

Definition. *The partial derivative of the function $z = f(x, y)$ with respect to x is the function of the two variables x and y appearing when $f(x, y)$ is differentiated with respect to x on condition that y is regarded as a constant.*

As has been said, the partial derivative of a function $z=f(x, y)$ with respect to x is designated by $f'_x(x, y)$; the symbols

$$\frac{\partial z}{\partial x}, z'_x, \frac{\partial f(x, y)}{\partial x}, \frac{\partial}{\partial x}[f(x, y)]$$

are also used for this purpose.

In contrast to the *ordinary* derivative we write ∂ instead of d and, besides the prime in the symbols z'_x and $f'_x(x, y)$, also write the subscript indicating the variable x with respect to which the differentiation is performed (by the way, the prime is often omitted, and we write z_x or $f_x(x, y)$).

It should be stressed that y is regarded constant *only* in the differentiation process with respect to x . After the expression for $f'_x(x, y)$ has been found both variables x and y can assume arbitrary admissible values. This exactly means that $f'_x(x, y)$ is a function of two independent variables. For instance, if $z=x^2y$ then $z'_x=2xy$.

It may of course happen that in a particular case $f'_x(x, y)$ depends solely on one variable and even is constant. For example, if $z=xy$ then $z'_x=y$, and if $z=2x+y^2$ then $z'_x=2$.

The *partial derivative of the function* $z=f(x, y)$ *with respect to* y is defined completely analogously:

$$f'_y(x, y) = \lim_{\Delta y \rightarrow 0} \frac{f(x, y+\Delta y) - f(x, y)}{\Delta y}$$

It is also denoted

$$\frac{\partial z}{\partial y}, z'_y, \frac{\partial f(x, y)}{\partial y} \text{ and } \frac{\partial}{\partial y}[f(x, y)]$$

A partial derivative being an *ordinary* derivative of a given function computed on condition that only one variable with respect to which the differentiation is performed changes, the computation of partial derivatives of elementary functions is carried out according to the well-known differentiation rules for elementary functions of one variable.

Examples.

(1) Let us find the derivatives z'_x and z'_y of the function $z=3axy-x^3-y^3$. Regarding y as a constant we obtain

$$\frac{\partial z}{\partial x} = 3ay - 3x^2$$

and regarding x as a constant we get

$$\frac{\partial z}{\partial y} = 3ax - 3y^2$$

(2) Let us find z'_x and z'_y for the function $z=x^y$ defined for $x>0$. In the differentiation with respect to x the magnitude z is

considered as a power function and in the differentiation with respect to y as an exponential function. Thus, we find

$$\frac{\partial z}{\partial x} = yx^{y-1}, \quad \frac{\partial z}{\partial y} = x^y \ln x$$

The particular values of the partial derivatives for given values $x = x_0$ and $y = y_0$ of the independent variables are denoted

$$\left(\frac{\partial z}{\partial x}\right)_{\substack{x=x_0 \\ y=y_0}}, \quad \left(\frac{\partial z}{\partial y}\right)_{\substack{x=x_0 \\ y=y_0}} \quad \text{or} \quad f'_x(x_0, y_0), \quad f'_y(x_0, y_0)$$

For example, if $z = 3axy - x^3 - y^3$ then

$$\left(\frac{\partial z}{\partial x}\right)_{\substack{x=a \\ y=0}} = -3a^2, \quad \left(\frac{\partial z}{\partial y}\right)_{\substack{x=a \\ y=0}} = 3a^2, \quad \left(\frac{\partial z}{\partial x}\right)_{\substack{x=a \\ y=a}} = \left(\frac{\partial z}{\partial y}\right)_{\substack{x=a \\ y=a}} = 0$$

Partial derivatives of a function of an arbitrary number of independent variables are defined analogously. In the case of n independent variables the notation of partial derivatives is the same as before; for instance,

$$\frac{\partial u}{\partial x_1} = f'_{x_1}(x_1, x_2, \dots, x_n) = \lim_{\Delta x_1 \rightarrow 0} \frac{f(x_1 + \Delta x_1, x_2, \dots, x_n) - f(x_1, x_2, \dots, x_n)}{\Delta x_1}$$

As an example, take the function $r = \sqrt{x^2 + y^2 + z^2}$. It expresses the distance between the point $P(x, y, z)$ and the origin. Computing its partial derivatives we obtain the expressions

$$r'_x = \frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2 + z^2}} = \frac{x}{r} = \cos \alpha$$

and

$$r'_y = \frac{\partial r}{\partial y} = \frac{y}{r} = \cos \beta, \quad r'_z = \frac{\partial r}{\partial z} = \frac{z}{r} = \cos \gamma$$

where $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ are the direction cosines of the polar radius of the point $P(x, y, z)$; these formulas will be used in what follows.

The absolute value of the partial derivative $z'_x = f'_x(x, y)$ or $z'_y = f'_y(x, y)$ is equal to the rate of change of the function $z = f(x, y)$ as only x or only y changes while the sign of f'_x or f'_y indicates the character of the variation (i.e. the increase or the decrease of the function).

The geometrical meaning of the partial derivatives of the function $z = f(x, y)$ is the following: $f'_x(x_0, y_0)$ is equal to the slope, relative to Ox , of the tangent line to the section of the surface $z = f(x, y)$ by the plane $y = y_0$ drawn through the point $M_0(x_0, y_0, z_0)$, that is, $f'_x(x_0, y_0) = \tan \alpha$ (Fig. 141a). (It is seen that $f'_x(x_0, y_0) < 0$ in the figure.)

The partial derivative $f'_y(x_0, y_0)$ is equal to the slope, relative to Oy , of the tangent line, at the point $M_0(x_0, y_0, z_0)$, to the section of the surface $z=f(x, y)$ by the plane $x=x_0$, i.e. $f'_y(x_0, y_0)=\tan\beta$ (Fig. 141b). (It is seen that $f'_y(x_0, y_0)>0$ in the figure.)

II. Partial differentials. The increment of the function $z=f(x, y)$ gained as only one of the variables changes is called a *partial increment of the function with respect to the corresponding variable*. The partial increments are denoted by the symbols

$$\Delta_x z = f(x + \Delta x, y) - f(x, y) \quad \text{and} \quad \Delta_y z = f(x, y + \Delta y) - f(x, y)$$

Definition. The *partial differential of the function $z=f(x, y)$ with respect to x* is the principal part of the partial increment $\Delta_x z = f(x + \Delta x, y) - f(x, y)$ proportional to the increment Δx of the independent variable x .

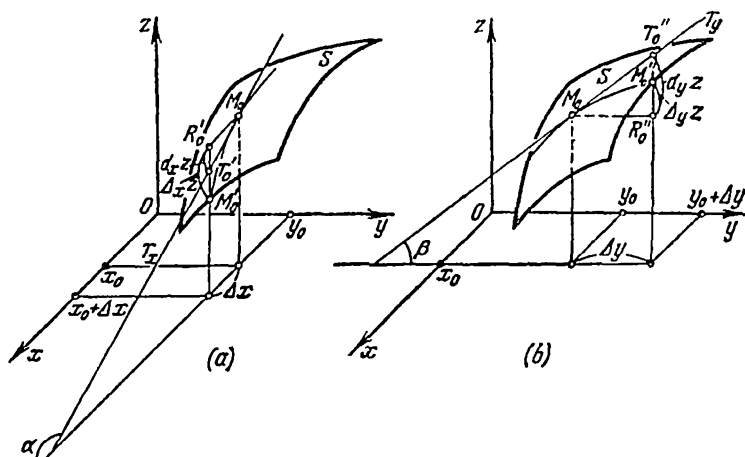


Fig. 141

The *partial differential with respect to y* is defined similarly.

The differentials of the independent variables x and y are simply understood as being equal to their increments:

$$dx = \Delta x, \quad dy = \Delta y$$

The partial differential with respect to x is denoted as $d_x z$ and with respect to y as $d_y z$.

As in the case of a function of one variable, it can readily be proved that if a function $z=f(x, y)$ possesses the partial differential with respect to x it also has the partial derivative z'_x and vice versa (see Sec. 50). Besides,

$$d_x z = \frac{\partial z}{\partial x} dx$$

Similarly, if a function $z=f(x, y)$ has the partial differential with respect to y it also has the partial derivative z'_y and vice versa, and

$$d_y z = \frac{\partial z}{\partial y} dy$$

Thus, a partial differential of a function of two independent variables is equal to the product of the partial derivative with respect to the corresponding variable by the differential of that variable.

The geometrical meaning of the partial increment $\Delta_x z$ is that it is equal to the increment of the z -coordinate of the point of the surface $z=f(x, y)$ as x receives the increment Δx (Fig. 141a). In the figure the increment $\Delta_x z$ is represented by the line segment $R'_0 M'_0$ and is negative: $\Delta_x z < 0$. The partial differential $d_x z$ is equal to the increment of the z -coordinate of the point of the tangent line T_x when the argument x gains the same increment. In the case shown in Fig. 141a we have $d_x z < 0$, the representing line segment being $R'_0 T'_0$.

Similarly, the partial increment $\Delta_y z$ is equal to the increment of the z -coordinate of the point of the surface as the argument y receives the increment Δy (Fig. 141b). In the figure the increment $\Delta_y z$ is positive ($\Delta_y z > 0$) and is represented by the line segment $R''_0 M''_0$. The partial differential $d_y z$ expresses the increment of the z -coordinate of the point of the tangent line T_y . In the case shown in Fig. 141b we have $d_y z > 0$, the representing segment being $R''_0 T''_0$.

From the formulas for the partial differentials we obtain

$$\frac{\partial z}{\partial x} = \frac{d_x z}{dx}, \quad \frac{\partial z}{\partial y} = \frac{d_y z}{dy}$$

It follows that partial derivatives, like ordinary derivatives, can be regarded as fractions whose numerators are the corresponding partial differentials and the denominators, the differentials of the corresponding independent variables. But the expressions $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are unified symbols which must not be understood as fractions, for, if we even agree that ∂x and ∂y denote dx and dy the symbols ∂z in the first and in the second cases designate different magnitudes (namely, $d_x z$ and $d_y z$).

As an instance, let us consider the *Mendeleev-Clapeyron*¹ equation

$$pv = RT$$

¹ Mendeleev, D. I. (1834-1907), the famous Russian scientist who constructed the periodic table of elements and discovered the periodic law.

Clapeyron, B.P.E (1799-1864), a French physicist.

and find $\frac{\partial p}{\partial v}$, $\frac{\partial v}{\partial T}$ and $\frac{\partial T}{\partial p}$ from it. We have

$$\frac{\partial p}{\partial v} = \frac{\partial}{\partial v} \left(\frac{RT}{v} \right) = -\frac{RT}{v^2}$$

$$\frac{\partial v}{\partial T} = \frac{\partial}{\partial T} \left(\frac{RT}{p} \right) = \frac{R}{p}$$

and

$$\frac{\partial T}{\partial p} = \frac{\partial}{\partial p} \left(\frac{pv}{R} \right) = \frac{v}{R}$$

Forming the product of these three partial derivatives we derive the relation

$$\frac{\partial p}{\partial v} \frac{\partial v}{\partial T} \frac{\partial T}{\partial p} = -\frac{RT}{v^2} \frac{R}{p} \frac{v}{R} = -\frac{RT}{pv} = -1$$

playing an important role in thermodynamics. (Note that if the symbols designating partial derivatives and involving ∂ were real fractions we would obtain 1 instead of -1 , which is incorrect.)

Partial increments and partial differentials of a function of any number of independent variables are defined by analogy with the case of a function of two variables. If $u = f(x_1, x_2, \dots, x_n)$ then, as above, we have

$$d_{x_1}u = \frac{\partial u}{\partial x_1} dx_1, \quad d_{x_2}u = \frac{\partial u}{\partial x_2} dx_2, \quad \dots,$$

that is, a *partial differential of a function of several independent variables with respect to one of them is equal to the product of the corresponding partial derivative by the differential of that variable.*

III. Total Differential.

1. **Total increment and their differential.** Let a function $z = f(x, y)$ be differentiable with respect to x and with respect to y . Then, using the partial derivatives of the function we can find arbitrarily accurate approximations to its increments corresponding to a variation of x for a constant value of y and to a variation of y for a constant x . Therefore it appears natural to try to find an approximation to the increment of the function $z = f(x, y)$ when both arguments simultaneously and independently gain arbitrary increments. In this general case the increment of the function is given by the formula

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y)$$

and is termed the *total increment*. In Fig. 142 we see that the total increment corresponding to the passage from the point $P(x, y)$ to the point $P_1(x + \Delta x, y + \Delta y)$ admits of the geometrical interpretation as the length of the line segment QM_1 .

The total increment of a function can be expressed by an extremely complicated formula involving Δx and Δy . But this relationship becomes quite simple when $f(x, y)$ is a linear function: $f(x, y) = ax + by + c$; in this case we readily find $\Delta z = a\Delta x + b\Delta y$.

It turns out that when $f(x, y)$ is not a linear function it is usually possible to choose, for given values of x and y , some constant coefficients A and B such that the expression $A\Delta x + B\Delta y$, which may not exactly coincide with Δz , differs from Δz by an

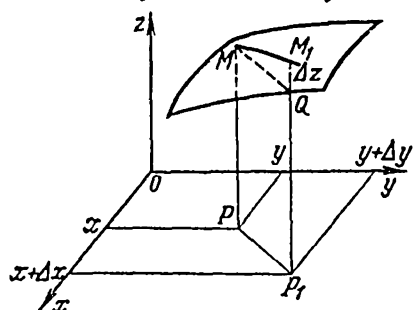


Fig. 142

infinitesimal of higher order than the distance ρ between the points P and P_1 as $\rho \rightarrow 0$ (we have $\rho = \sqrt{\Delta x^2 + \Delta y^2}$):

$$\Delta z = A\Delta x + B\Delta y + o(\rho)$$

In this case the sum $A\Delta x + B\Delta y$ is called the *total differential* of the function $z = f(x, y)$ and is denoted dz or $df(x, y)$:

$$dz = A\,dx + B\,dy \quad (*)$$

(as before, $\Delta x = dx$ and $\Delta y = dy$).

Thus, as $\rho \rightarrow 0$, the difference between the total increment Δz of the function $z = f(x, y)$ and its total differential dz is an infinitesimal of higher order than ρ . This is briefly stated thus: dz is the *principal part* of the increment Δz . Now we can state the definition:

Definition. A function $z = f(x, y)$ is said to be *differentiable at a given point* (x, y) if its total increment is representable in the form

$$\Delta z = A\Delta x + B\Delta y + o(\rho), \quad \rho = \sqrt{\Delta x^2 + \Delta y^2}$$

where A and B are independent of Δx and Δy .

The principal part of the total increment of a differentiable function which is a linear function of the increments of the independent variables is called the *total differential* (or, briefly, the differential) of the function:

$$dz = A\Delta x + B\Delta y$$

A function differentiable at each point of a domain is said to be *differentiable in that domain*.

It is clear that if a function is differentiable at a given point it is continuous at it. But the fact that a function is continuous at a point does not imply that it is differentiable at that point. Moreover, in the case of a function of one variable the existence of the differential is equivalent to the existence of the derivative (see Sec. 52), but this is no longer true for a function of two

variables: a function $z = f(x, y)$ possessing both partial derivatives $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ may have no total differential, that is, is not necessarily differentiable! At the end of the present section we shall prove that the condition that *both partial derivatives of the function $z = f(x, y)$ not only exist at the given point but also exist in a neighbourhood of the point and are continuous in it* is sufficient for the function to be differentiable at that point.

Now, assuming that the function in question possesses a total differential, we shall prove a proposition showing how the total differential can be found:

Theorem. The total differential of a function of two independent variables is equal to the sum of the products of the partial derivatives of the function by the differentials of the corresponding independent variables.

Proof. Formula (*) expressing the differential is valid for arbitrary dx and dy and hence, in particular, for $dy = 0$. In the latter case the total increment Δz coincides with the partial increment $\Delta_x z$, and we obtain

$$d_x z = A dx$$

whence (see Sec. 110, II)

$$A = \frac{d_x z}{dx} = f'_x(x, y)$$

The fact that $B = f'_y(x, y)$ is proved similarly. Consequently, the expression for the total differential for given values of x and y is written

$$dz = f'_x(x, y) dx + f'_y(x, y) dy$$

or

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

which is what we set out to prove.

Since

$$\frac{\partial z}{\partial x} dx = d_x z \quad \text{and} \quad \frac{\partial z}{\partial y} dy = d_y z$$

we obtain

$$dz = d_x z + d_y z$$

that is, the differential of a function of two independent variables is equal to the sum of its partial differentials.

Examples. (1) Let $z = 3axy - x^3 - y^3$. Then

$$dz = (3ay - 3x^2) dx + (3ax - 3y^2) dy.$$

(2) For the function $z = x^y$ ($x > 0$) we have

$$dz = yx^{y-1} dx + x^y \ln x dy$$

In both examples the partial derivatives are continuous and functions are therefore differentiable.

The definition of the total differential is extended without changes to functions of any number of independent variables. If $u = f(x_1, x_2, \dots, x_n)$ then the total increment of the function is

$$\Delta u = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) - f(x_1, x_2, \dots, x_n)$$

If the total increment is expressible as the sum of the principal part (linear with respect to the increments of the independent variables) and an infinitesimal of higher order the function is differentiable and its total differential, i.e. the principal (linear) part of its total increment, is

$$du = \frac{\partial u}{\partial x_1} dx_1 + \frac{\partial u}{\partial x_2} dx_2 + \dots + \frac{\partial u}{\partial x_n} dx_n$$

The relationship between the increment Δu and the differential du is given by the formula $\Delta u = du + o(\rho)$ where $o(\rho)$ is an infinitesimal of higher order relative to

$$\rho = \sqrt{\Delta x_1^2 + \Delta x_2^2 + \dots + \Delta x_n^2}$$

As in the case of functions of two variables, the existence of continuous partial derivatives with respect to all independent variables guarantees the existence of the total differential.

Note. Let $u = f(x_1, x_2, \dots, x_n)$ be a function of n independent variables differentiable in a given domain and let its total differential be identically equal to zero for all the values of the independent variables and arbitrary $dx_1, dx_2, \dots, dx_n$:

$$du = \frac{\partial u}{\partial x_1} dx_1 + \frac{\partial u}{\partial x_2} dx_2 + \dots + \frac{\partial u}{\partial x_n} dx_n \equiv 0$$

Then, putting $dx_1 \neq 0, dx_2 = dx_3 = \dots = dx_n = 0$ we conclude that $\frac{\partial u}{\partial x_1} \equiv 0$. Similarly, all the other partial derivatives $\frac{\partial u}{\partial x_2}, \frac{\partial u}{\partial x_3}, \dots, \frac{\partial u}{\partial x_n}$ also turn out to be identically equal to zero. Consequently, in this case u is independent of $x_1, x_2, \dots, x_n$ and is of constant magnitude.

II*. Sufficient condition for differentiability. We shall prove a theorem providing a sufficient condition for differentiability of a function of two variables (it was already mentioned after the definition of differentiability).

Theorem. If a function $z = f(x, y)$ possesses partial derivatives in a neighbourhood of a point $P(x, y)$ and these derivatives $f'_x(x, y)$ and $f'_y(x, y)$ are continuous at that point the function is differentiable at the point $P(x, y)$.

Proof. Let us consider the total increment of the function $z = f(x, y)$:

$$\Delta z = f(x + \Delta x, y + \Delta y) - f(x, y)$$

On adding and subtracting the expression $f(x, y + \Delta y)$ we rewrite this formula as

$$\Delta z = [f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)] + [f(x, y + \Delta y) - f(x, y)]$$

In other words, the increment of the function corresponding to the passage from the point P to the point P_1 (see Fig. 143) is representable as the sum of two increments, namely, of the increment gained by the function for the given constant value of x (that is, corresponding to the passage from the point P to the point N) and the increment for the given constant value $y + \Delta y$ of the other argument (which corresponds to the passage from the point N to the point P_1).¹

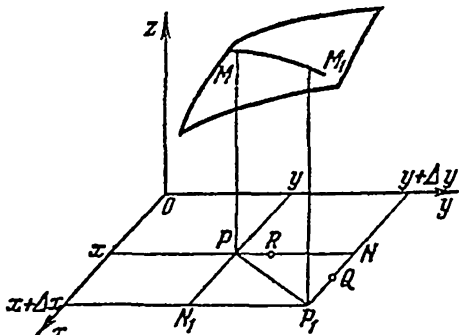


Fig. 143

The expression in the first parentheses is the increment of the function $f(x, y)$ for the constant value $y + \Delta y$ of the ordinate received when x is given the increment Δx . To this increment we can apply Lagrange's formula of finite increments (see Sec. 57), which results in

$$f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y) = f'_x(\xi_1, y + \Delta y) \Delta x$$

(the position of the point $Q(\xi_1, y + \Delta y)$ with coordinates ξ_1 and $y + \Delta y$, $\xi_1 \in [x, x + \Delta x]$ is shown in Fig. 143).

Similarly, applying Lagrange's formula to the expression in the second parentheses as increment of the function corresponding to the variation of only one argument y we obtain

$$f(x, y + \Delta y) - f(x, y) = f'_y(x, \xi_2) \Delta y$$

(the point $R(x, \xi_2)$ with coordinates x and ξ_2 , $\xi_2 \in [y, y + \Delta y]$, lies on the line segment PN).

Thus, we have

$$\Delta z = f'_x(\xi_1, y + \Delta y) \Delta x + f'_y(x, \xi_2) \Delta y$$

By the hypothesis, the derivatives f'_x and f'_y are continuous functions. Therefore, since the points Q and R approach the point P and $P_1 \rightarrow P$, we can write

$$f'_x(\xi_1, y + \Delta y) = f'_x(x, y) + \varepsilon_1 \text{ and } f'_y(x, \xi_2) = f'_y(x, y) + \varepsilon_2$$

¹ The consideration is similar if we take the increments of the function corresponding to the paths PN_1 and N_1P_1 .

where e_1 and e_2 tend to zero together with Δx and Δy , and hence together with $\rho = PP_1 = \sqrt{\Delta x^2 + \Delta y^2}$. On substituting these expressions into the above formula of Δz we receive

$$\Delta z = [f'_x(x, y) + e_1] \Delta x + [f'_y(x, y) + e_2] \Delta y$$

that is,

$$\Delta z = f'_x(x, y) \Delta x + f'_y(x, y) \Delta y + \alpha \quad (*)$$

where $\alpha = e_1 \Delta x + e_2 \Delta y$. By virtue of the inequalities $|\Delta x| \leq \rho$ and $|\Delta y| \leq \rho$ we have

$$|\alpha| = |e_1 \Delta x + e_2 \Delta y| \leq |e_1| |\Delta x| + |e_2| |\Delta y| \leq \rho (|e_1| + |e_2|)$$

Consequently,

$$\frac{|\alpha|}{\rho} \leq |e_1| + |e_2|$$

Since e_1 and e_2 tend to zero as $\rho \rightarrow 0$ the ratio $\frac{|\alpha|}{\rho}$ also tends to zero and therefore α is an infinitesimal of higher order than ρ . It follows that, according to the definition, the sum of the first two terms on the right-hand side of equality (*) which is linear in Δx and Δy is the total differential of the function at the point $P(x, y)$.

This theorem is readily extended to the case of more than two independent variables.

112. Geometrical Interpretation of the Total Differential of a Function of Two Variables. As is known, the derivative and the differential of a function of one variable are connected with the tangent line to the graph of the function. It turns out that, similarly, the partial derivatives and the total differential of a function of two variables are connected with the *tangent plane* to the surface serving as the graph of the function.

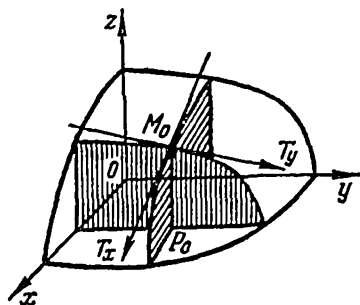


Fig. 144

Let $z = f(x, y)$ be a differentiable function at a point (x_0, y_0) . Consider the sections of the surface S representing this function by the planes

$y = y_0$ and $x = x_0$. Let us draw the tangent lines $M_0 T_x$ and $M_0 T_y$ at the point $M_0(x_0, y_0, z_0)$ (see Fig. 144) to the plane curves obtained in the sections. The plane T passing through these lines which meet at the point M_0 is called the *tangent plane* to the surface S at the point M_0 , the point M_0 being the *point of tangency* (the *point of contact*) of the plane T and the surface S .

Let us derive the equation of the tangent plane. The straight line $M_0 T_x$ is in the plane $y = y_0$ parallel to the plane Oxz , its

slope relative to the x -axis being equal to $f'_x(x_0, y_0)$. Therefore the line M_0T_x is specified by the equations

$$z - z_0 = f'_x(x_0, y_0)(x - x_0), \quad y = y_0$$

The equations of the straight line M_0T_y are found similarly:

$$z - z_0 = f'_y(x_0, y_0)(y - y_0), \quad x = x_0$$

The plane T passing through the point $M_0(x_0, y_0, z_0)$, its equation is written as

$$z - z_0 = A(x - x_0) + B(y - y_0)$$

The straight lines M_0T_x and M_0T_y are in the plane T , and therefore the coordinates of the points of these lines satisfy the equation of the plane. Substituting the expressions of $z - z_0$ and $y - y_0$ from the equation of the straight line M_0T_x into the equation of the plane we obtain

$$f'_x(x_0, y_0)(x - x_0) = A(x - x_0)$$

whence

$$A = f'_x(x_0, y_0)$$

Furthermore, we similarly find that

$$B = f'_y(x_0, y_0)$$

Thus, the equation of the tangent plane takes the form

$$z - z_0 = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0) \quad (*)$$

Later on (see Sec. 119) we shall prove that the tangent line drawn through the point $M_0(x_0, y_0, z_0)$ to any curve lying on the surface S and passing through that point is in the tangent plane (*).

The geometrical meaning of the total differential of a function of two independent variables is clarified by the following proposition:

Theorem. The total differential of a function $z = f(x, y)$ is equal to the increment of the z -coordinate of the tangent plane drawn to the graph of the function at the corresponding point.

Proof. The right-hand side of the equation of the tangent plane (*) coincides with the expression of the total differential of the function $z = f(x, y)$ (since $x - x_0 = dx$, $y - y_0 = dy$). It follows that the equation of the tangent plane can be written in the form

$$z - z_0 = (dz)_{\substack{x=x_0 \\ y=y_0}}$$

where z_0 is the z -coordinate of the point of tangency, z is the coordinate of the moving point of the tangent plane and $(dz)_{\substack{x=x_0 \\ y=y_0}}$ the total differential of the function $z = f(x, y)$ computed for $x = x_0$ and $y = y_0$.

If x receives an increment Δx and y an increment Δy (Fig. 145) the function gains the corresponding increment Δz represented by the line segment R_1M_1 , equal to the increment of the z -coordinate of the point of the surface S while the differential dz is represented by the segment R_1T_1 which is the increment of the z -coordinate of the corresponding point of the tangent plane T .

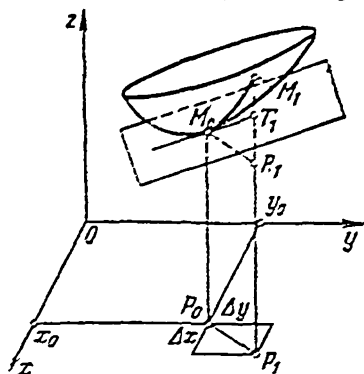


Fig. 145

113. Applying Total Differentials to Approximate Calculations. In § 53 we discussed the application of the differential of a function of several variables to approximate calculations. In this section we shall treat analogous problems for functions of several variables and their differentials. For the sake of simplicity we shall take functions of two variables although all the facts we discuss are readily extended to functions of any number of variables.

The fundamental idea of the application of the differential to approximate calculations lies in the *replacement of the total increment of a function by its total differential*, that is, it reduces the approximate relation

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) \approx dz$$

valid for small Δx and Δy . Replacing dz by its expression in terms of the derivatives we write this approximate equality

$$f(x_0 + \Delta x, y_0 + \Delta y) \approx f(x_0, y_0) + f'_x(x_0, y_0)\Delta x + f'_y(x_0, y_0)\Delta y$$

If we put $x_0 + \Delta x = x$, $y_0 + \Delta y = y$ this approximate relation is brought to the form

$$f(x, y) \approx f(x_0, y_0) + f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0)$$

The latter formula indicates that the replacement of the increment of the function by its total differential is equivalent to the replacement, in the vicinity of the point (x_0, y_0) , of the function $f(x, y)$ by the linear function written on the right-hand side of formula (*). This means, geometrically, that a portion of the surface $z = f(x, y)$ is replaced by the corresponding part of the tangent plane to the surface at the point $M_0(x_0, y_0, z_0)$.

Formula (*) enables us to compute approximations to the values of the function $f(x, y)$ when x is close to x_0 and y is close to y_0 after the values $f(x_0, y_0)$, $f'_x(x_0, y_0)$ and $f'_y(x_0, y_0)$ have been determined.

Example. Let us write down the approximate formula for computing the values of the function

$$z = \ln(xy + 2y^2 - 2x)$$

in the vicinity of the point $(1, 1)$. Here we have $x_0 = 1$, $y_0 = 1$, and the computations yield

$$\begin{aligned}\frac{\partial z}{\partial x} &= \frac{1}{xy + 2y^2 - 2x} (y - 2), & \left(\frac{\partial z}{\partial x}\right)_{\substack{x=1 \\ y=1}} &= -1, \\ \frac{\partial z}{\partial y} &= \frac{1}{xy + 2y^2 - 2x} (x + 4y), & \left(\frac{\partial z}{\partial y}\right)_{\substack{x=1 \\ y=1}} &= 5\end{aligned}$$

Thus, since $z_0 = 0$, we obtain

$$\ln(xy + 2y^2 - 2x) \approx -(x - 1) + 5(y - 1)$$

Let the reader compute several values of this function by means of the approximate formula and compare them with the tabular values.

Now we shall show how the limiting absolute error in the determination of a value of a function of two variables is found when the limiting absolute errors of its arguments are known. The reader should recall the subject matter of Sec. 53 (whose notation we retain here) before proceeding to what follows.

Suppose it is required to compute the value of a given function $z = f(x, y)$ corresponding to given values of the arguments x and y which are only known approximately:

$$x = x_0 + dx, \quad |dx| < e_x; \quad y = y_0 + dy, \quad |dy| < e_y$$

where x_0 and y_0 are the given approximate values of the arguments and e_x and e_y the corresponding limiting absolute errors. To find the limiting absolute error e_z of the value of the function it is necessary to estimate the modulus of the difference between the exact value $f(x, y)$ and the approximate value $f(x_0, y_0)$. Formula (*) implies

$$\begin{aligned}|f(x, y) - f(x_0, y_0)| &\approx |f'_x(x_0, y_0)dx + f'_y(x_0, y_0)dy| \leq \\ &\leq |f'_x(x_0, y_0)||dx| + |f'_y(x_0, y_0)||dy| < \\ &< |f'_x(x_0, y_0)|e_x + |f'_y(x_0, y_0)|e_y\end{aligned}$$

Hence, we can put

$$e_z = |f'_x(x_0, y_0)|e_x + |f'_y(x_0, y_0)|e_y \quad (**)$$

On dividing e_z by $f(x_0, y_0)$ we get the limiting relative error δ_z .

As an example, let us determine the limiting relative errors of a product and of a quotient.

(1) Let $z = xy$. If x_0 and y_0 are approximate values of the arguments, formula (**) yields

$$e_z = |y_0| e_x + |x_0| e_y$$

On dividing by $|z_0| = |x_0 y_0|$ we find δ_z :

$$\delta_z = \frac{e_x}{|x_0|} + \frac{e_y}{|y_0|} = \delta_x + \delta_y$$

Thus, *the limiting relative error in the computation of a product is equal to the sum of the limiting relative errors of the factors.* Of course, this rule holds for any number of factors.

(2) Let $z = \frac{y}{x}$ and $z_0 = \frac{y_0}{x_0}$. Then

$$e_z = \left| -\frac{y_0}{x_0^2} \right| e_x + \left| \frac{1}{x_0} \right| e_y$$

and

$$\delta_z = \frac{e_x}{|x_0|} + \frac{e_y}{|y_0|} = \delta_x + \delta_y$$

that is, *the limiting relative error of a quotient is equal to the sum of the limiting relative errors of the dividend and the divisor.* (Thus, the limiting relative errors appearing in the computations of the product and of the quotient of two magnitudes coincide.)

Formula (**) also enables us to solve the reverse problem: *given an admissible error e_z of the determination of the function $f(x, y)$ and some approximate values x_0 and y_0 of the arguments, it is required to find the admissible values of e_x and e_y .* The conditions of the problem do not specify e_x and e_y uniquely since these two magnitudes are only connected by one relation (**). Therefore we can choose the values of e_x and e_y depending on which what can be decreased more easily, that is, which of the magnitudes x and y can be measured with a greater accuracy. If it turns out that the given approximate values x_0 and y_0 are not measured sufficiently accurately they should be measured again with the required accuracy, after which the new approximate value of the function is computed (see Example 2 on page 167).

114. Derivatives and Differentials of Higher Orders.

I. Suppose that a function $z = f(x, y)$ possesses the partial derivatives

$$\frac{\partial z}{\partial x} = f'_x(x, y), \quad \frac{\partial z}{\partial y} = f'_y(x, y)$$

These derivatives, in their turn, are functions of x and y . The partial derivatives of these functions (provided they exist) are called the *second partial derivatives* or the *partial derivatives of the second order* of the given function $f(x, y)$. Each of the derivatives $f''_x(x, y)$

and $f'_y(x, y)$ (referred to as the *first derivatives* or the *derivatives of the first order*) has two partial derivatives, and thus we obtain four partial derivatives of the second order denoted

$$\begin{aligned}\frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right) &= \frac{\partial^2 z}{\partial x^2} = f''_{xx} = z''_{xx}, & \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right) &= \frac{\partial^2 z}{\partial x \partial y} = f''_{xy} = z''_{xy} \\ \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right) &= \frac{\partial^2 z}{\partial y \partial x} = f''_{yx} = z''_{yx}, & \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial y} \right) &= \frac{\partial^2 z}{\partial y^2} = f''_{yy} = z''_{yy}\end{aligned}$$

The derivatives f''_{xy} and f''_{yx} are termed *mixed*; the former is obtained when the function is first differentiated with respect to x and then with respect to y and the latter when the indicated order of the differentiations is reversed.

For instance, taking the function $z = x^3y^2 - 3xy^3 - xy + 1$ we receive

$$\begin{aligned}\frac{\partial z}{\partial x} &= 3x^2y^2 - 3y^3 - y, & \frac{\partial z}{\partial y} &= 2x^3y - 9xy^2 - x \\ \frac{\partial^2 z}{\partial x^2} &= 6xy^2, & \frac{\partial^2 z}{\partial y^2} &= 2x^3 - 18xy \\ \frac{\partial^2 z}{\partial x \partial y} &= 6x^2y - 9y^2 - 1, & \frac{\partial^2 z}{\partial y \partial x} &= 6x^2y - 9y^2 - 1\end{aligned}$$

In this example the mixed derivatives turn out to be identically coincident. The following theorem indicates that this fact is a particular case of a general rule which we state without proof:

Theorem. If the mixed partial derivatives of the second order of a function $z = f(x, y)$ are continuous they coincide:

$$f''_{xy}(x, y) = f''_{yx}(x, y)$$

The condition of continuity of the mixed derivatives is essential: if it is violated the derivatives are not necessarily equal.

Thus, under the above conditions a function $z = f(x, y)$ of two variables possesses not four but three different partial derivatives of the second order:

$$\frac{\partial^2 z}{\partial x^2}, \quad \frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x} \quad \text{and} \quad \frac{\partial^2 z}{\partial y^2}$$

The partial derivatives of the partial derivatives of the second order are called *partial derivatives of the third order* (or *third partial derivatives*), etc.

The above theorem on the equality of the mixed derivatives of the second order implies the following general proposition:

The result of a repeated differentiation of a function of two independent variables does not depend on the order of the differentiations if all the partial derivatives in question are continuous.

For example, let us prove that

$$\frac{\partial^3 z}{\partial x \partial y^2} = \frac{\partial^3 z}{\partial y \partial x^2}$$

Using the theorem on the equality of the mixed derivatives of the second order we write

$$\frac{\partial^2 y}{\partial x \partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial^2 z}{\partial x \partial y} \right) = \frac{\partial}{\partial y} \left(\frac{\partial^2 z}{\partial y \partial x} \right) = \frac{\partial^2 z}{\partial y \partial x \partial y}$$

which is what we wanted to prove. For all the other cases this general proposition is proved analogously.

It is readily seen that the function $z = f(x, y)$ has $n+1$ partial derivatives of the n th order; they can be denoted as

$$\frac{\partial^n z}{\partial x^n}, \quad \frac{\partial^n z}{\partial x^{n-1} \partial y}, \quad \frac{\partial^n z}{\partial x^{n-2} \partial y^2}, \quad \dots, \quad \frac{\partial^n z}{\partial x^2 \partial y^{n-2}}, \quad \frac{\partial^n z}{\partial x \partial y^{n-1}}, \quad \frac{\partial^n z}{\partial y^n}$$

As a rule, an elementary function of two independent variables possesses partial derivatives of any order throughout its domain of definition (except, possibly, at separate points or lines).

Partial derivatives of higher orders of functions of any number of independent variables are defined similarly. The theorem on the independence of the result of repeated differentiations of the order of the differentiations applies to this case as well. For example, if $u = f(x, y, z)$ then

$$\frac{\partial^3 u}{\partial x \partial y \partial z} = \frac{\partial^3 u}{\partial x \partial z \partial y} = \frac{\partial^3 u}{\partial y \partial x \partial z} = \frac{\partial^3 u}{\partial y \partial z \partial x} = \frac{\partial^3 u}{\partial z \partial x \partial y} = \frac{\partial^3 u}{\partial z \partial y \partial x}$$

Repeated differentiation of functions of several variables is performed practically by finding the required derivatives of higher order one after another.

II. The total differential of function $z = f(x, y)$ (also termed the *first differential* or the *differential of the first order*) is expressed by the formula

$$dz = f'_x(x, y) dx + f'_y(x, y) dy = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

and is dependent on the arguments x and y and also on their differentials dx and dy . The differentials dx , dy do not depend on x and y and are regarded as constant magnitudes when total differentials of higher order are computed in succession (although they can take on arbitrary values; see the analogous remark for functions of one argument in Sec. 55).

The total differential of the second order (the second differential) is the total differential of the total differential of the first order dz computed under the assumption that dx and dy retain (arbitrary) constant values. If all the partial derivatives of the second order are continuous this definition implies the formula

$$d^2 z = d \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right) = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right) dx + \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right) dy$$

for the partial differential of the second order. On computing the

partial derivatives of the expressions in the parentheses we obtain

$$\frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right) = \frac{\partial^2 z}{\partial x^2} dx + \frac{\partial^2 z}{\partial x \partial y} dy$$

and

$$\frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \right) = \frac{\partial^2 z}{\partial x \partial y} dx + \frac{\partial^2 z}{\partial y^2} dy$$

The substitution of these expressions into the above formula of d^2z yields, after similar terms have been combined, the final formula

$$d^2z = \frac{\partial^2 z}{\partial x^2} dx^2 + 2 \frac{\partial^2 z}{\partial x \partial y} dx dy + \frac{\partial^2 z}{\partial y^2} dy^2$$

Total differentials of higher order are defined similarly by applying, in succession, the above procedure the required number of times on condition that all the partial derivatives involved are continuous. As an instance, let the reader verify that the total differential of the third order is given by the formula

$$d^3z = \frac{\partial^3 z}{\partial x^3} dx^3 + 3 \frac{\partial^3 z}{\partial x^2 \partial y} dx^2 dy + 3 \frac{\partial^3 z}{\partial x \partial y^2} dx dy^2 + \frac{\partial^3 z}{\partial y^3} dy^3$$

and think about the general form of the subsequent total differentials.

115. Reconstructing a Function from Its Total Differential.

I. The case of two independent variables. Let x and y be independent variables and $P(x, y)$ and $Q(x, y)$ functions of these variables continuous together with their partial derivatives in a domain D (bounded or unbounded). This condition is supposed to hold throughout this section and will not be stipulated in the statements of the theorems.

We say that an expression $P(x, y)dx + Q(x, y)dy$ is a total differential if there is a function $u(x, y)$ whose total differential coincides with this expression:

$$du = P(x, y)dx + Q(x, y)dy \quad (*)$$

Let us analyse the conditions guaranteeing that a given expression $P(x, y)dx + Q(x, y)dy$ is a total differential. Remember that for a function of one variable a product of the form $f(x)dx$ where $f(x)$ is a continuous function is always a total differential since there exists a function y such that $dy = f(x)dx$. Indeed, any anti-derivative of $f(x)$ can be taken as the function y : $y = \int f(x)dx$. It turns out that for functions of two variables the situation is essentially different: an expression $P(x, y)dx + Q(x, y)dy$ is by far not always a total differential, and for a function $u(x, y)$ such that $du = P(x, y)dx + Q(x, y)dy$ to exist some additional conditions should hold. We shall give these conditions in the two theorems

below: the first of them expresses a necessary condition for an expression $P dx + Q dy$ to be a total differential and the second provides a sufficient condition.

Theorem (Necessary Condition). If an expression

$$P(x, y) dx + Q(x, y) dy$$

is a total differential in a domain D then at all the points of the domain D there holds the equality

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (**)$$

Proof. By the hypothesis, there exists a function $u(x, y)$ for which equality (*) is fulfilled. Since

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy$$

equality (*) implies

$$P(x, y) = \frac{\partial u}{\partial x} \quad \text{and} \quad Q(x, y) = \frac{\partial u}{\partial y}$$

On differentiating the first of these relations with respect to y and the second with respect to x we obtain

$$\frac{\partial P}{\partial y} = \frac{\partial^2 u}{\partial x \partial y}, \quad \frac{\partial Q}{\partial x} = \frac{\partial^2 u}{\partial y \partial x}$$

The mixed partial derivatives of the second order of the function $u(x, y)$ which are the first derivatives of the functions $P(x, y)$ and $Q(x, y)$ being continuous, the theorem in Sec. 114 implies that they are equal, that is,

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

which is what we set out to prove.

The theorem proved shows that if condition (**) does not hold the expression $P dx + Q dy$ is not a total differential. For example, in the expression $y dx - x dy$ we have $P = y$ and $Q = -x$. Therefore $P'_y = 1$, $Q'_x = -1$; these derivatives are continuous throughout the xy -plane but condition (**) is violated: $1 \neq -1$. Hence, there exists no function whose total differential coincides with the expression $y dx - x dy$.

In the formulation of the necessary condition the domain D can be quite arbitrary but in the sufficient condition stated below it is essential that the domain D is *simply connected* (see Sec. 107, III).

Theorem (Sufficient Condition). If the condition

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (**)$$

is fulfilled at all the points of a simply connected domain D there exists a function $u(x, y)$ defined in this domain whose total differential is

$$du = P(x, y) dx + Q(x, y) dy \quad (*)$$

It can be shown that there exists an infinitude of functions having a given total differential: they are all given by the expression $u(x, y) + C$ where C is an arbitrary constant and $u(x, y)$ is any function whose total differential is equal to the given expression $P dx + Q dy$ (this proposition is similar to the analogous proposition for functions of one variable stated in Sec. 78, and we leave its proof to the reader; to this end, it is advisable to make use of the remark on page 384).

Here we shall prove this theorem for the particular case of a simply connected domain D shaped as a rectangle with sides parallel to the coordinate axes (Fig. 146). The rectangular domain D can be finite or infinite; this means that it can be a strip, or a quadrant or a half-plane bounded by straight lines parallel to the coordinate axes; it can also coincide with the whole plane. These cases are obtained in the limit when some of the sides of the finite rectangle shown in Fig. 146 "recede" to infinity. For the general case of an arbitrary simply connected domain this theorem will be proved in Sec. 141.

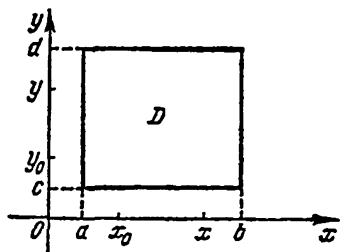


Fig. 146

As was shown, the function $u(x, y)$ (if it exists) must satisfy the two conditions

$$(1) \quad \frac{\partial u}{\partial x} = P(x, y) \quad \text{and} \quad (2) \quad \frac{\partial u}{\partial y} = Q(x, y)$$

There are infinitely many functions satisfying the first of these conditions: they are all given by the formula

$$u(x, y) = \int_{x_0}^x P(x, y) dx + \varphi(y) \quad (***)$$

where $[x_0, x] \subset (a, b)$ and $\varphi(y)$ is an arbitrary function of y . For, the differentiation of expression (***) with respect to x immediately results in $u'_x = P(x, y)$.

Now let us choose a function $\varphi(y)$ for which the second condition is fulfilled. To this end, we differentiate the expression of $u(x, y)$ with respect to y and equate the result to $Q(x, y)$:

$$\frac{\partial u}{\partial y} = \frac{\partial}{\partial y} \int_{x_0}^x P(x, y) dx + \varphi'(y) = Q(x, y)$$

It follows that

$$\varphi'(y) = Q(x, y) - \frac{\partial}{\partial y} \int_{x_0}^x P(x, y) dx$$

It can be shown (but we shall not present the proof here) that if the functions $P(x, y)$ and $P'_y(x, y)$ are continuous the operations of differentiation with respect to y and integration with respect to x can be interchanged (this is known as *Leibniz' rule*). Therefore, taking into account equality (**), we get

$$\begin{aligned} \frac{\partial}{\partial y} \int_{x_0}^x P(x, y) dx &= \int_{x_0}^x \frac{\partial P(x, y)}{\partial y} dx = \int_{x_0}^x \frac{\partial Q(x, y)}{\partial x} dx = \\ &= Q(x, y) - Q(x_0, y) \end{aligned}$$

The substitution of the value of the integral $\int_{x_0}^x \frac{\partial Q(x, y)}{\partial x} dx$ thus found into the formula expressing $\varphi'(y)$ yields

$$\varphi'(y) = Q(x, y) - Q(x, y) + Q(x_0, y) = Q(x_0, y)$$

The summands involving x have thus cancelled out, and on the right-hand side we have the function $Q(x_0, y)$ dependent solely on y . On integrating, we receive

$$\varphi(y) = \int_{y_0}^y Q(x_0, y) dy + C$$

where $[y_0, y] \subset (c, d)$ and C is an arbitrary constant. Thus, the function $\varphi(y)$ has been determined; its substitution into formula (***) gives us the final result:

$$u(x, y) = \int_{x_0}^x P(x, y) dx + \int_{y_0}^y Q(x_0, y) dy + C \quad (A)$$

The order of these operations can of course be reversed, that is, we can start with finding a function $u(x, y)$ satisfying the second condition (its expression involves an arbitrary function of the argument x) and then determine this function by using the first condition. Let the reader check that this leads to the equivalent formula

$$u(x, y) = \int_{y_0}^y Q(x, y) dy + \int_{x_0}^x P(x, y_0) dx + C \quad (B)$$

The initial point (x_0, y_0) entering into both formulas (A) and (B) can be chosen quite arbitrarily within the domain D . It is

advisable to choose it in such a way that the expression for $Q(x_0, y)$ (or for $P(x, y_0)$) becomes as simple as possible.

The proof of the theorem for the case of a rectangular domain with sides parallel to the coordinate axes given here incidentally provides formulas (A) and (B) for reconstructing a function $u(x, y)$ from its total differential.

In practical determination of a function $u(x, y)$ from its total differential formulas (A) and (B) are rarely used since it is more convenient to apply the modified computational scheme involving indefinite integrals, which is demonstrated by the example below.

Consider the expression

$$(e^y + x) dx + (xe^y - 2y) dy$$

Let us check that it is a total differential and determine the function $u(x, y)$. Here we have

$$P = e^y + x, \quad Q = xe^y - 2y, \quad \frac{\partial P}{\partial y} = e^y, \quad \frac{\partial Q}{\partial x} = e^y$$

These functions are everywhere continuous and $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, that is the sufficient condition for the existence of the function $u(x, y)$ is fulfilled. Now let us find this function. We have the two relations

$$(1) \quad \frac{\partial u}{\partial x} = e^y + x \quad \text{and} \quad (2) \quad \frac{\partial u}{\partial y} = xe^y - 2y$$

The indefinite integration of (1) results in

$$u(x, y) = \int (e^y + x) dx + \varphi(y) = xe^y + \frac{x^2}{2} + \varphi(y)$$

Furthermore,

$$\frac{\partial u}{\partial y} = xe^y + \varphi'(y) = xe^y - 2y$$

whence we obtain $\varphi'(y) = -2y$ since the terms with x mutually cancel out. Hence,

$$\varphi(y) = -y^2 + C$$

Consequently,

$$u(x, y) = xe^y + \frac{x^2}{2} - y^2 + C$$

The correctness of the result can be checked directly.

II. The case of three independent variables. The method applied above for finding a function of two independent variables from its total differential is generalized almost without changes to the case of three arguments. Therefore we shall limit ourselves to a brief discussion.

Let $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ be functions of independent variables x, y, z continuous together with their partial derivatives in a spatial domain Ω . There holds the following theorem:

Theorem. If the expression

$$P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz \quad (*)$$

is a total differential of a function $u(x, y, z)$ in the domain Ω then at all the points of the domain Ω there must hold the conditions

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z} \quad (**)$$

Indeed, since the function $u(x, y, z)$ whose total differential is expression (*) exists, we have

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz$$

This yields the three equalities

$$\frac{\partial u}{\partial x} = P, \quad \frac{\partial u}{\partial y} = Q, \quad \frac{\partial u}{\partial z} = R \quad (***)$$

which immediately imply identities (**) in accordance with the equality of the corresponding mixed partial derivatives of the function $u(x, y, z)$. For instance,

$$\frac{\partial Q}{\partial z} = \frac{\partial^2 u}{\partial y \partial z} \quad \text{and} \quad \frac{\partial R}{\partial y} = \frac{\partial^2 u}{\partial z \partial y}, \quad \text{i.e.} \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \quad \text{etc.}$$

Now let the domain Ω be a rectangular parallelepiped with faces parallel to the coordinate planes (or a domain obtained as some faces of such a parallelepiped recede to infinity). Then, if conditions (**) hold throughout the domain Ω there exists a function $u(x, y, z)$ whose total differential is equal to the expression (*).

The argument similar to that in the case of a plane region leads us to the following formula:

$$u(x, y, z) = \int_{x_0}^x P(x, y, z) dx + \int_{y_0}^y Q(x_0, y, z) dy + \int_{z_0}^z R(x_0, y_0, z) dz + C$$

It is of course supposed that the points (x_0, y_0, z_0) and (x, y, z) belong to the domain Ω (Fig. 147). The order of operations in the determination of the function u can be of course changed; this leads to the corresponding modifications of the final formula.

A more general class of domains Ω for which conditions (**) imply the existence of the function $u(x, y, z)$ will be discussed in Sec. 143.

We shall consider an example demonstrating the reconstruction of a function $u(x, y, z)$ from its total differential without using the formula derived above. Let us find the general expression for

the function $u(x, y, z)$ if

$$du = (2xyz + \ln y) dx + \left(x^2z + \frac{x}{y}\right) dy + (x^2y - 2z) dz$$

Here we have

$$P = 2xyz + \ln y, \quad Q = x^2z + \frac{x}{y}, \\ R = x^2y - 2z$$

and the partial derivatives of these functions are continuous in the half-space $y > 0$. The expression of du is in fact a total differential since

$$\frac{\partial P}{\partial y} = 2xz + \frac{1}{y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = x^2 = \frac{\partial R}{\partial y}, \\ \frac{\partial R}{\partial x} = 2xy = \frac{\partial P}{\partial z}$$

Now let us find the expression of the function $u(x, y, z)$. We begin with the third condition (***):

$$\frac{\partial u}{\partial z} = x^2y - 2z$$

The general form of the function u satisfying this condition is

$$u = \int (x^2y - 2z) dz + \varphi(x, y) = x^2yz - z^2 + \varphi(x, y)$$

where $\varphi(x, y)$ is an arbitrary function of x and y which should now be chosen in such a way that the first two conditions (***) hold. To this end, we must have

$$\frac{\partial u}{\partial x} = 2xyz + \frac{\partial \varphi}{\partial x} = 2xyz + \ln y$$

and

$$\frac{\partial u}{\partial y} = x^2z + \frac{\partial \varphi}{\partial y} = x^2z + \frac{x}{y}$$

whence $\frac{\partial \varphi}{\partial x} = \ln y$ and $\frac{\partial \varphi}{\partial y} = \frac{x}{y}$, that is,

$$d\varphi = \ln y dx + \frac{x}{y} dy$$

Since $\frac{\partial(\ln y)}{\partial y} = \frac{\partial\left(\frac{x}{y}\right)}{\partial x} = \frac{1}{y}$, we have arrived at the problem of finding a function $\varphi(x, y)$ of two variables having a given total differential. Therefore we use the procedure described in I to obtain

$$\varphi(x, y) = \int \ln y dx = x \ln y + \psi(y)$$

and

$$\frac{\partial \varphi}{\partial y} = \frac{x}{y} + \psi'(y) = \frac{x}{y}$$

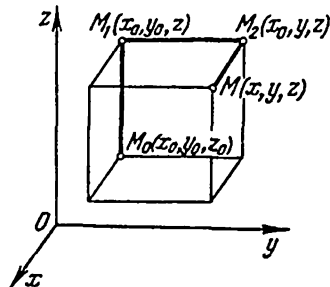


Fig. 147

It is obvious that $\psi'(y)=0$, i.e. $\psi(y)=C$. Finally, we get

$$u = x^2 y z - z^2 + x \ln y + C$$

Let the reader check directly the correctness of this result.

116. Differentiating Composite Functions. Rules for Finding Differentials of Functions.

I. Differentiating composite functions. Consider a function $z=f(u, v)$ possessing the continuous partial derivatives z'_u and z'_v . By Sec. 111, this function is differentiable and its increment Δz is representable in the form

$$\Delta z = \frac{\partial z}{\partial u} \Delta u + \frac{\partial z}{\partial v} \Delta v + \alpha \quad (*)$$

where α is an infinitesimal of higher order than $\sqrt{\Delta u^2 + \Delta v^2}$. Now suppose that u and v are, in their turn, differentiable functions of an independent variable x , i.e.

$$u = \varphi(x) \quad \text{and} \quad v = \psi(x)$$

Thus,

$$z = f[\varphi(x), \psi(x)] = F(x)$$

which means that z is a *composite function* of one variable x . Our aim is to express the derivative $\frac{dz}{dx}$ in terms of the *partial derivatives* $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ and the derivatives of the functions u and v with respect to x . To this end, we give the argument x an increment Δx . This makes u and v receive, respectively, some increments Δu and Δv in terms of which Δz is expressed by formula (*). Let us divide both sides of this formula by Δx :

$$\frac{\Delta z}{\Delta x} = \frac{\partial z}{\partial u} \frac{\Delta u}{\Delta x} + \frac{\partial z}{\partial v} \frac{\Delta v}{\Delta x} + \frac{\alpha}{\Delta x}$$

The next step is to pass to the limit as $\Delta x \rightarrow 0$. By the hypothesis, we have $\lim_{\Delta x \rightarrow 0} \frac{\Delta u}{\Delta x} = \frac{du}{dx}$ and $\lim_{\Delta x \rightarrow 0} \frac{\Delta v}{\Delta x} = \frac{dv}{dx}$. The ratio $\frac{\alpha}{\Delta x}$ can be represented as

$$\frac{\alpha}{\Delta x} = \frac{\alpha}{\sqrt{\Delta u^2 + \Delta v^2}} \frac{\sqrt{\Delta u^2 + \Delta v^2}}{\Delta x} = \frac{\alpha}{\sqrt{\Delta u^2 + \Delta v^2}} \sqrt{\left(\frac{\Delta u}{\Delta x}\right)^2 + \left(\frac{\Delta v}{\Delta x}\right)^2}$$

The first factor on the right-hand side tends to zero in accordance with the definition of α while the second factor tends to a definite number, namely, to $\sqrt{\left(\frac{du}{dx}\right)^2 + \left(\frac{dv}{dx}\right)^2}$. Consequently, $\lim_{\Delta x \rightarrow 0} \frac{\alpha}{\Delta x} = 0$.

Finally, noting that the values of $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ depend *solely* on the

chosen value of x determining the values of u and v and are independent of Δx we obtain

$$\frac{dz}{dx} = \frac{\partial z}{\partial u} \frac{du}{dx} + \frac{\partial z}{\partial v} \frac{dv}{dx}$$

Clearly, this formula is a generalization of the differentiation rule for a composite function of one variable.

Now let us suppose that z is a composite function of two independent variables x and y , that is,

$$z = f(u, v)$$

where

$$u = \varphi(x, y) \quad \text{and} \quad v = \psi(x, y)$$

Thus,

$$z = f[\varphi(x, y), \psi(x, y)] = F(x, y)$$

We also suppose that all the functions involved possess continuous partial derivatives and therefore are *differentiable*.

To find z'_x we must consider y constant, and then u and v become functions of only one variable x ; therefore we arrive at the case treated above. The only distinction is that the ordinary derivatives $\frac{du}{dx}$ and $\frac{dv}{dx}$ are now replaced by the partial derivatives $\frac{\partial u}{\partial x}$ and $\frac{\partial v}{\partial x}$, and we thus obtain

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x}$$

and, similarly,

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y}$$

Hence,

A partial derivative of a composite function is equal to the sum of the products of the derivatives of the given function with respect to the intermediate arguments (i.e. u and v) by the partial derivatives of these arguments with respect to the corresponding independent variable (x or y).

Example. Let us take the function $z = e^{xy} \sin(x+y)$. If $xy = u$ and $x+y = v$, that is $z = e^u \sin v$, we obtain

$$\begin{aligned} z'_x &= \frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x} = e^u \sin v \cdot y + e^u \cos v \cdot 1 = \\ &= e^{xy} [y \sin(x+y) + \cos(x+y)] \end{aligned}$$

and

$$\begin{aligned} z'_y &= \frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y} = e^u \sin v \cdot x + e^u \cos v \cdot 1 = \\ &= e^{xy} [x \sin(x+y) + \cos(x+y)] \end{aligned}$$

where r is the polar radius of the point (x, y, z) in space: $r = \sqrt{x^2 + y^2 + z^2}$. We have

$$\frac{\partial u}{\partial x} = \frac{\partial \left(\frac{1}{r} \right)}{\partial x} = \frac{d \left(\frac{1}{r} \right)}{dr} \frac{\partial r}{\partial x} = -\frac{1}{r^2} \frac{x}{r} = -\frac{x}{r^3}$$

$$\frac{\partial u}{\partial y} = -\frac{y}{r^3}, \quad \frac{\partial u}{\partial z} = -\frac{z}{r^3}$$

and, furthermore,

$$\frac{\partial^2 u}{\partial x^2} = -\frac{r^3 - 3r^2 \frac{\partial r}{\partial x} x}{r^6} = -\frac{r^3 - 3rx^2}{r^6} = -\frac{r^2 - 3x^2}{r^6}$$

$$\frac{\partial^2 u}{\partial y^2} = -\frac{r^2 - 3y^2}{r^6}, \quad \frac{\partial^2 u}{\partial z^2} = -\frac{r^2 - 3z^2}{r^6}$$

Note that the function $u = \frac{1}{r}$ satisfies the relation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

(2) Let $z = x^3 e^{u^2}$ where $u = \varphi(x)$. Then

$$\frac{\partial z}{\partial x} = 3x^2 e^{u^2} \quad \text{and} \quad \frac{dz}{dx} = 3x^2 e^{u^2} + x^3 e^{u^2} 2u\varphi'(x)$$

II. Invariance of the form of the first differential and the general rules for computing differentials.

Theorem. The total differential of a function $z = f(u, v)$ retains its form irrespective of whether its arguments u and v are independent variables or functions of some other independent variables.

If u and v are independent variables then the total differential is

$$dz = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv$$

Now let u and v be functions of independent variables x and y possessing continuous partial derivatives. The formulas for $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ derived in I indicate that these derivatives are also continuous; consequently, z as function of x and y is differentiable, and the definition of the total differential implies

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

On replacing $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ by the expressions given in I we get

$$dz = \left(\frac{\partial z}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial x} \right) dx + \left(\frac{\partial z}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial z}{\partial v} \frac{\partial v}{\partial y} \right) dy$$

Now, regrouping the terms, we arrive at the relation

$$dz = \frac{\partial z}{\partial u} \left(\frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy \right) + \frac{\partial z}{\partial v} \left(\frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy \right)$$

where the expressions in the parentheses are, respectively, equal to du and dv . Consequently,

$$dz = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv$$

which is what we intended to prove.

Thus, a function of two variables, like a function of one variable, possesses the invariance property of the form of its first differential (see Sec. 51, III). Total differentials of higher order do not possess this property.

Analogously, if $z = f(u, v, \dots, w)$ then we always have

$$dz = \frac{\partial z}{\partial u} du + \frac{\partial z}{\partial v} dv + \dots + \frac{\partial z}{\partial w} dw$$

irrespective of whether the arguments $u, v, \dots, w$ depend or do not depend on some other independent variables.

In the computation of the differential of a function of several independent variables we can use the same simple rules as in the case of one independent variable (Sec. 51). Let $u, v, \dots, w$ be functions of any number of independent variables. Then the following rules analogous to those in Sec. 51 hold:

- (1) $d(u + v + \dots + w) = du + dv + \dots + dw$;
- (2) $d(uv) = v du + u dv$ and, in particular, $d(Cu) = C du$ where $C = \text{const}$;
- (3) $d\left(\frac{u}{v}\right) = \frac{v du - u dv}{v^2}$

The proof of these rules follows directly from the invariance of the form of the first differential. As an instance, let us prove the second rule. The form of the differential being independent of the nature of the arguments, we can suppose that they are independent variables. Then

$$d(uv) = \frac{\partial}{\partial u}(uv) du + \frac{\partial}{\partial v}(uv) dv = v du + u dv$$

which is what we wanted to prove.

The other rules are proved similarly. These rules sometimes enable us to simplify the calculations.

Example. Let us find $d \arctan \frac{y}{x}$. Putting $\frac{y}{x} = u$ we obtain

$$d \arctan \frac{y}{x} = (\arctan u)' du = \frac{1}{1 + \frac{y^2}{x^2}} d\left(\frac{y}{x}\right)$$

Now, using the third rule, we finally get

$$d \arctan \frac{y}{x} = \frac{1}{1 + \frac{y^2}{x^2}} \frac{x dy - y dx}{x^2} = -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy$$

By the way, it follows immediately that

$$\frac{\partial}{\partial x} \arctan \frac{y}{x} = -\frac{y}{x^2 + y^2} \quad \text{and} \quad \frac{\partial}{\partial y} \arctan \frac{y}{x} = \frac{x}{x^2 + y^2}$$

117. Existence Theorem for an Implicit Function. In our course we already dealt with implicit functions, that is, with those defined by means of equations connecting variable magnitudes. We remind the reader that an implicit function of one variable is specified by an equation

$$F(x, y) = 0 \quad (*)$$

As was noted in Sec. 12, there are cases when equation (*) does not determine a function; for instance, the equation $x^2 + y^2 + 5 = 0$ has no real roots and hence y cannot be regarded as a function of x . In the general case, for a more complicated equation, it is extremely difficult to find out for what values of x equation (*) has roots and what is their number; then the question as to whether the equation specifies a function remains open.

Here we are going to establish certain conditions guaranteeing that one of the variables involved in an equation $F(x, y) = 0$ is specified as a function of the other. These conditions are given by the following theorem (which we present without proof):

Theorem (On the Existence of an Implicit Function). Let $F(x, y)$ be a function continuous together with its partial derivatives in a neighbourhood of a point $M_0(x_0, y_0)$. If

$$F(x_0, y_0) = 0 \quad \text{and} \quad F'_y(x_0, y_0) \neq 0$$

then for the values of x lying sufficiently close to x_0 the equation

$$F(x, y) = 0$$

possesses a unique solution $y = \varphi(x)$ depending continuously on x such that $\varphi(x_0) = y_0$. Besides, the function $\varphi(x)$ also possesses a continuous derivative.

Below are examples demonstrating the application of this theorem.

(1) Let $F(x, y) = x^2 + y^2 - R^2$. The equation

$$x^2 + y^2 - R^2 = 0$$

specifies a circle. At any point $M_0(x_0, y_0)$ on this circle such that $y_0 \neq 0$ (see Fig. 148) all the conditions of the above theorem are fulfilled:

$$x_0^2 + y_0^2 - R^2 = 0 \quad \text{and} \quad F'_y(x_0, y_0) = 2y_0 \neq 0$$

The function $y = \sqrt{R^2 - x^2}$ (the sign of the root is chosen in accordance with the sign of y_0) is the only continuous function satisfying the given equation and the condition $\sqrt{R^2 - x_0^2} = y_0$. Taking the opposite sign of the root we obtain another continuous function (satisfying the same equation) which is no longer equal to y_0 at the point $x = x_0$ (it assumes the value $-y_0$ for $x = x_0$). It should also be noted that

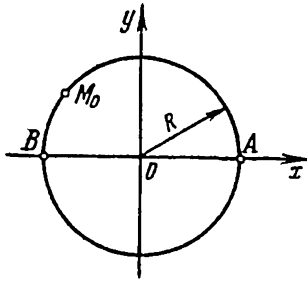


Fig. 148

if for different values of x we choose the corresponding values of y with different signs the resultant function will be discontinuous (cf. footnote on page 37) while the theorem speaks of the existence of a unique continuous function.

The situation is quite different in the vicinity of the points $A(R, 0)$ and $B(-R, 0)$. For the point A there are no real values of y corresponding to the values $x > R$, and for $x \leq R$ both conti-

nuous real solutions $y = +\sqrt{R^2 - x^2}$ and $y = -\sqrt{R^2 - x^2}$ turn into zero for $x = R$. The matter is that here the partial derivative F'_y turns into zero: $F'_y(R, 0) = 0$, and therefore the conditions of the existence theorem are violated. On the other hand, we have $F'_x(R, 0) = 2R \neq 0$ at this point, and hence x can be expressed as a uniquely determined continuous function of y , which is again clearly seen in Fig. 148.

(2) Let an implicit function be specified by the equation

$$x^3y + \ln y - x = 0 \quad (*)$$

In this case $F(x, y) = x^3y + \ln y - x$. At the point $M_0(1, 1)$ we have $F(1, 1) = 0$. The partial derivatives $F'_x = 3x^2y - 1$ and $F'_y = x^3 + \frac{1}{y}$ are continuous in the vicinity of this point, and $F'_y(1, 1) = 2 \neq 0$. Hence, by the existence theorem, there is a unique function $y = \varphi(x)$ satisfying the given equation such that $\varphi(1) = 1$. Although we have established the existence of the function $\varphi(x)$, it cannot be expressed as an *elementary* function of x since the equation is not solvable algebraically in y . Particular values of the function $\varphi(x)$ can be found by setting certain values of x and applying the methods of Sec. 77 to solving the resultant equations. The function $y = \varphi(x)$ being continuous, for the values of x lying close to 1, the corresponding equation possesses a root close to 1.

An implicit function of two variables is specified by an equation of the form

$$F(x, y, z) = 0$$

connecting the three variables x , y and z . The question of the existence of such a function is elucidated by the following theorem.

Theorem. Let $F(x, y, z)$ be a function continuous together with its partial derivatives in a neighbourhood of a point $M_0(x_0, y_0, z_0)$. If

$$F(x_0, y_0, z_0) = 0 \quad \text{and} \quad F'_z(x_0, y_0, z_0) \neq 0$$

then the equation $F(x, y, z) = 0$ possesses a unique solution $z = \varphi(x, y)$, defined in a sufficiently small neighbourhood of the point (x_0, y_0) , depending continuously on x and y and satisfying the condition $\varphi(x_0, y_0) = z_0$. Moreover, the function $\varphi(x, y)$ also has continuous partial derivatives.

It should also be noted that if $F'_z = 0$ at the point M_0 but some other derivative, say F'_y , is different from zero ($F'_y(x_0, y_0, z_0) \neq 0$) the equation may not define z as a function of x and y but specifies y as a function of x and z .

118. Differentiating Implicit Functions. Suppose that the conditions of the existence theorem for an implicit function stated in the foregoing section are fulfilled for a function $F(x, y)$ so that the equation

$$F(x, y) = 0$$

specifies y as function of x : $y = \varphi(x)$. The substitution of the function $\varphi(x)$ for y into this equation leads to the identity

$$F[x, \varphi(x)] \equiv 0$$

It follows that the derivative of the function $F(x, y)$ (where $y = \varphi(x)$) with respect to x is also identically zero. On differentiating this expression according to the differentiation rule for a composite function (see Sec. 116) we find

$$F'_x + F'_y \frac{dy}{dx} = 0$$

whence

$$y' = \frac{dy}{dx} = -\frac{F'_x}{F'_y}$$

This formula expresses the derivative of the implicit function $y = \varphi(x)$ in terms of the partial derivatives of the given function $F(x, y)$.¹ The derivative y' does not exist at the point (x, y) for which $F'_y = 0$. But, as was mentioned, for such values of x and y the existence of the function $\varphi(x)$ itself is not guaranteed.

¹ Another method for computing the derivative of an implicit function of one variable was discussed in Sec. 45, II.

Examples. (1) Let $F(x, y) \equiv x^2 + y^2 - 1 = 0$. We have $\frac{\partial F}{\partial x} = 2x$ and $\frac{\partial F}{\partial y} = 2y$, and hence $y' = -\frac{2x}{2y} = -\frac{x}{y}$.

(2) Taking the equation $F(x, y) \equiv x^3y + \ln y - x = 0$ we find, according to the solution of Example 2 in Sec. 117, that $y' = -\frac{3x^2y-1}{x^3+\frac{1}{y}}$.

Now let us consider an equation

$$F(x, y, z) = 0$$

specifying z as a function $z = \varphi(x, y)$ of the independent variables x and y . The substitution of $\varphi(x, y)$ for z into the equation results in the identity

$$F[x, y, \varphi(x, y)] \equiv 0$$

Consequently, the partial derivatives of the function $F(x, y, z)$ (with z replaced by $\varphi(x, y)$) with respect to x and y must also be equal to zero. On differentiating, we obtain

$$F'_x + F'_z \frac{\partial z}{\partial x} = 0 \quad \text{and} \quad F'_y + F'_z \frac{\partial z}{\partial y} = 0$$

whence

$$\frac{\partial z}{\partial x} = -\frac{F'_x}{F'_z}, \quad \frac{\partial z}{\partial y} = -\frac{F'_y}{F'_z}$$

These formulas express the partial derivatives of the implicit function $z = \varphi(x, y)$ in terms of the derivatives of the given function $F(x, y, z)$. If $F'_z = 0$ the formulas become senseless.

Examples. (1) Let us find the partial derivatives of the function z specified by the equation $F(x, y, z) \equiv x^2 + y^2 + z^2 - 1 = 0$ (the equation of unit sphere). We have $\frac{\partial F}{\partial x} = 2x$, $\frac{\partial F}{\partial y} = 2y$ and $\frac{\partial F}{\partial z} = 2z$; hence,

$$\frac{\partial z}{\partial x} = -\frac{x}{z}, \quad \frac{\partial z}{\partial y} = -\frac{y}{z}$$

In this example it is possible to find the explicit expression of the function and to check the correctness of the results obtained.

(2) Take the equation $F(x, y, z) \equiv e^{-xy} - 2z + e^z = 0$. The differentiation yields

$$\frac{\partial F}{\partial x} = -ye^{-xy}, \quad \frac{\partial F}{\partial y} = -xe^{-xy}, \quad \frac{\partial F}{\partial z} = -2 + e^z$$

Therefore,

$$\frac{\partial z}{\partial x} = \frac{ye^{-xy}}{e^z - 2}, \quad \frac{\partial z}{\partial y} = \frac{xe^{-xy}}{e^z - 2}$$

In this case it is impossible to express z as an explicit function, and the values of the partial derivatives should be found for given values of the variables x and y after the corresponding value of the function z has been computed in some way.

In the general case an equation of the form

$$F(x, y, z, \dots, u) = 0$$

specifies u as a function of $x, y, z, \dots$; by analogy with the foregoing cases, we find

$$\frac{\partial u}{\partial x} = -\frac{F'_x}{F'_u}, \quad \frac{\partial u}{\partial y} = -\frac{F'_y}{F'_u}, \quad \frac{\partial u}{\partial z} = -\frac{F'_z}{F'_u}, \quad \dots$$

§ 3. Applications of Differential Calculus to Geometry

119. Surfaces. If a surface in space is specified by an equation

$$z = f(x, y)$$

the equation of the tangent plane to the surface at its point $M_0(x_0, y_0, z_0)$ is written (see Sec. 112) as

$$z - z_0 = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0)$$

Let us consider a surface S specified in the coordinate space $Oxyz$ by an equation of the general form

$$F(x, y, z) = 0$$

not solved with respect to a coordinate. Suppose that in a neighbourhood of a point $M_0(x_0, y_0, z_0)$ the function $F(x, y, z)$ satisfies the conditions of the existence theorem for an implicit function so that the equation $F(x, y, z) = 0$ determines z as a function of x and y : $z = f(x, y)$ and $f(x_0, y_0) = z_0$. According to Sec. 118, the derivatives of this function at the point (x_0, y_0) are expressed by the formulas

$$f'_x(x_0, y_0) = -\frac{F'_x(x_0, y_0, z_0)}{F'_z(x_0, y_0, z_0)}, \quad f'_y(x_0, y_0) = -\frac{F'_y(x_0, y_0, z_0)}{F'_z(x_0, y_0, z_0)}$$

The equation of the tangent plane can be rewritten as

$$z - z_0 = -\frac{F'_x(x_0, y_0, z_0)}{F'_z(x_0, y_0, z_0)}(x - x_0) - \frac{F'_y(x_0, y_0, z_0)}{F'_z(x_0, y_0, z_0)}(y - y_0)$$

or

$$F'_x(x_0, y_0, z_0)(x - x_0) + F'_y(x_0, y_0, z_0)(y - y_0) + F'_z(x_0, y_0, z_0)(z - z_0) = 0$$

Using the subscript "0" to designate that the partial derivatives are taken at the point of tangency $M_0(x_0, y_0, z_0)$, that is, for $x=x_0$, $y=y_0$ and $z=z_0$, we can write this equation in the abbreviated form

$$\left(\frac{\partial F}{\partial x}\right)_0(x-x_0) + \left(\frac{\partial F}{\partial y}\right)_0(y-y_0) + \left(\frac{\partial F}{\partial z}\right)_0(z-z_0) = 0$$

This form of the equation of the tangent plane also remains valid when $\left(\frac{\partial F}{\partial z}\right)_0 = 0$. The equality $\left(\frac{\partial F}{\partial z}\right)_0 = 0$ simply means that the tangent plane at the given point of the surface is parallel to the z -axis. The equation of the tangent plane becomes senseless if all the three partial derivatives of the function $F(x, y, z)$ turn into zero at the point M_0 : $\left(\frac{\partial F}{\partial x}\right)_0 = \left(\frac{\partial F}{\partial y}\right)_0 = \left(\frac{\partial F}{\partial z}\right)_0 = 0$. Points of this kind are called *singular*, and we shall not treat them in our course.

Definition. The straight line perpendicular to the tangent plane at the point of tangency is called the *normal to the surface* at that point.

The equation of the tangent plane to a surface $F(x, y, z) = 0$ at its point $M_0(x_0, y_0, z_0)$ being known, we can write the equations of the normal to the surface at that point on the basis of the rules established in analytic geometry. These equations have the form

$$\frac{x-x_0}{\left(\frac{\partial F}{\partial x}\right)_0} = \frac{y-y_0}{\left(\frac{\partial F}{\partial y}\right)_0} = \frac{z-z_0}{\left(\frac{\partial F}{\partial z}\right)_0}$$

If a surface is represented by an equation $z=f(x, y)$ the equations of the normal to the surface at its point $M_0(x_0, y_0, z_0)$ are written as

$$\frac{x-x_0}{f'_x(x_0, y_0)} = \frac{y-y_0}{f'_y(x_0, y_0)} = \frac{z-z_0}{-1}$$

Using the equation of the tangent plane we can readily find the expressions for the direction cosines of the normal; let the reader determine the direction cosines for the cases when a surface is specified by an equation explicitly and implicitly (these formulas will be used in what follows).

Examples. (1) Let us derive the equation of the tangent plane and the equations of the normal to the sphere

$$x^2 + y^2 + z^2 = R^2$$

at its point $M_0(x_0, y_0, z_0)$. We have

$$\left(\frac{\partial F}{\partial x}\right)_0 = 2x_0, \quad \left(\frac{\partial F}{\partial y}\right)_0 = 2y_0, \quad \left(\frac{\partial F}{\partial z}\right)_0 = 2z_0$$

and hence the equation of the tangent plane is

$$x_0(x - x_0) + y_0(y - y_0) + z_0(z - z_0) = 0$$

or

$$x_0x + y_0y + z_0z = x_0^2 + y_0^2 + z_0^2 = R^2$$

The equations of the normal have the form

$$\frac{x - x_0}{x_0} = \frac{y - y_0}{y_0} = \frac{z - z_0}{z_0} \quad \text{or} \quad \frac{x}{x_0} = \frac{y}{y_0} = \frac{z}{z_0}$$

It follows that the normal to that sphere passes through the origin.

(2) Let us form the equation of the tangent plane, parallel to the plane $x + y - z = 0$, to the ellipsoid of revolution $x^2 + \frac{y^2}{2} + z^2 = 1$.

Since $\left(\frac{\partial F}{\partial x}\right)_0 = 2x_0$, $\left(\frac{\partial F}{\partial y}\right)_0 = y_0$ and $\left(\frac{\partial F}{\partial z}\right)_0 = 2z_0$ the equation of the tangent plane at the corresponding point $M_0(x_0, y_0, z_0)$ (whose location is yet unknown) is written

$$2x_0(x - x_0) + y_0(y - y_0) + 2z_0(z - z_0) = 0 \quad (*)$$

The parallelism condition for the planes yields the relations

$$\frac{2x_0}{1} = \frac{y_0}{1} = \frac{2z_0}{-1}$$

To these relations we should add the condition expressing the fact that the point M_0 lies on the ellipsoid:

$$x_0^2 + \frac{y_0^2}{2} + z_0^2 = 1$$

We thus arrive at the system of three equations with three unknowns. Its solutions are the coordinates of the two points $M'_0\left(\frac{1}{2}, 1, -\frac{1}{2}\right)$ and $M''_0\left(-\frac{1}{2}, -1, \frac{1}{2}\right)$. On substituting the values of the coordinates thus found into equation (*) we obtain the equations of the two planes satisfying the conditions of the problem:

$$x + y - z = 2 \quad \text{and} \quad x + y - z = -2$$

In Sec. 112 the tangent plane to a surface was defined as the one passing through the tangent lines to the two plane curves obtained in the sections of the surface by the two coordinate planes Oxz and Oyz . Now we shall establish a property of the tangent plane which, by the way, accounts for its name.

If a curve L lies entirely on a surface the tangent line to that curve at its point M_0 is in the tangent plane to the surface drawn

through this point. In other words, if all the curves lying on the surface and passing through the point M_0 are drawn the tangent lines to these curves at the point M_0 are all in a common plane which is the tangent plane to the surface.

Let a surface be specified by an equation $F(x, y, z) = 0$ and let L be a smooth curve on that surface passing through a point M_0 of the surface and represented by parametric equations $x = x(t)$, $y = y(t)$, $z = z(t)$, the value of the parameter t corresponding to the point M_0 being equal to t_0 . The condition that the curve L lies entirely on the surface is expressed by the relation

$$F[x(t), y(t), z(t)] = 0$$

This relation being an identity for all the values of t , we can differentiate it with respect to t in accordance with the differentiation rule for a composite function, which results in

$$\frac{\partial F}{\partial x} x' + \frac{\partial F}{\partial y} y' + \frac{\partial F}{\partial z} z' = 0$$

Indicating the values of the derivatives at the point M_0 by the subscript "0" we can write down the latter relation in the form

$$\left(\frac{\partial F}{\partial x}\right)_0 x'_0 + \left(\frac{\partial F}{\partial y}\right)_0 y'_0 + \left(\frac{\partial F}{\partial z}\right)_0 z'_0 = 0 \quad (**)$$

The values $\left(\frac{\partial F}{\partial x}\right)_0$, $\left(\frac{\partial F}{\partial y}\right)_0$, $\left(\frac{\partial F}{\partial z}\right)_0$ of the partial derivatives depend solely on the equation of the surface and on the chosen point M_0 while the values of the derivatives x'_0 , y'_0 , z'_0 additionally depend on the choice of the curve L and are equal to the projections of the direction vector of the tangent line to the curve L at the point M_0 (see Sec. 68). Equality (**) indicates that any such *tangent vector* is perpendicular to one and the same *normal* $\{(F'_x)_0, (F'_y)_0, (F'_z)_0\}$ (which is a direction vector of the normal to the surface). (Equality (**) expresses the fact that the scalar product of these vectors is equal to zero.) This exactly means that all the tangent lines to the curves lying on the surface and passing through the point M_0 are in one plane which is the tangent plane to the surface.

120. Space Curve as Line of Intersection of Two Surfaces. In Sec. 68 we studied parametric equations of space curves. A curve in space can also be represented as a line of intersection of two surfaces:

$$F(x, y, z) = 0, \quad \Phi(x, y, z) = 0 \quad (*)$$

It is clear, geometrically, that the tangent to this curve at a point $M_0(x_0, y_0, z_0)$ is the line of intersection of the tangent planes to these surfaces drawn through that point. Forming the equations

of these planes

$$\left(\frac{\partial F}{\partial x}\right)_0 (x-x_0) + \left(\frac{\partial F}{\partial y}\right)_0 (y-y_0) + \left(\frac{\partial F}{\partial z}\right)_0 (z-z_0) = 0$$

$$\left(\frac{\partial \Phi}{\partial x}\right)_0 (x-x_0) + \left(\frac{\partial \Phi}{\partial y}\right)_0 (y-y_0) + \left(\frac{\partial \Phi}{\partial z}\right)_0 (z-z_0) = 0$$

we arrive at the equations of the sought-for tangent as line of intersection of the two planes. The direction vector $\mathbf{T}$ of this tangent line can be found as the vector product of the normal vectors of the two tangent planes (Fig. 149). These normal vectors are

$$\mathbf{N}_1 = \left\{ \left(\frac{\partial F}{\partial x}\right)_0, \left(\frac{\partial F}{\partial y}\right)_0, \left(\frac{\partial F}{\partial z}\right)_0 \right\} \text{ and } \mathbf{N}_2 = \left\{ \left(\frac{\partial \Phi}{\partial x}\right)_0, \left(\frac{\partial \Phi}{\partial y}\right)_0, \left(\frac{\partial \Phi}{\partial z}\right)_0 \right\}$$

whence

$$\mathbf{T} = \mathbf{N}_1 \times \mathbf{N}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \left(\frac{\partial F}{\partial x}\right)_0 & \left(\frac{\partial F}{\partial y}\right)_0 & \left(\frac{\partial F}{\partial z}\right)_0 \\ \left(\frac{\partial \Phi}{\partial x}\right)_0 & \left(\frac{\partial \Phi}{\partial y}\right)_0 & \left(\frac{\partial \Phi}{\partial z}\right)_0 \end{vmatrix}$$

Knowing the direction vector $\mathbf{T}$ and the point of tangency M_0 we can write down the canonical equations of the tangent line and also form the equation of the normal plane. Here we do not write these equations in the general form and proceed to the example below demonstrating the course of the calculations.

Example. Let a curve be the line of intersection of the two cylinders:

$$x^2 + y^2 = 10$$

and

$$y^2 + z^2 = 25$$

We shall form the equations of the tangent line and of the normal plane to this curve at the point $(1, 3, 4)$.

To begin with, let us determine the projections of the vectors $\mathbf{N}_1$ and $\mathbf{N}_2$:

$$\left(\frac{\partial F}{\partial x}\right)_0 = 2, \quad \left(\frac{\partial F}{\partial y}\right)_0 = 6, \quad \left(\frac{\partial F}{\partial z}\right)_0 = 0$$

and

$$\left(\frac{\partial \Phi}{\partial x}\right)_0 = 0, \quad \left(\frac{\partial \Phi}{\partial y}\right)_0 = 6, \quad \left(\frac{\partial \Phi}{\partial z}\right)_0 = 8$$

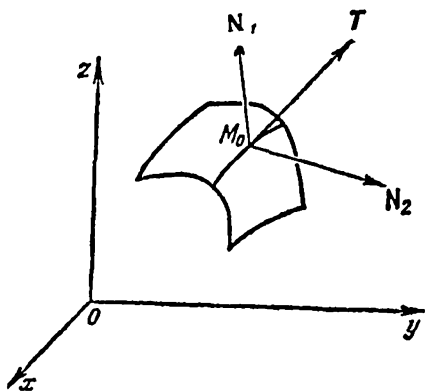


Fig. 149

Next we compute their vector product:

$$\mathbf{T} = \mathbf{N}_1 \times \mathbf{N}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 6 & 0 \\ 0 & 6 & 8 \end{vmatrix} = 48\mathbf{i} - 16\mathbf{j} + 12\mathbf{k}$$

Now, knowing the vector $\mathbf{T}$ and the point of tangency we write down the sought-for equations: the equations of the tangent line are

$$\frac{x-1}{12} = \frac{y-3}{-4} = \frac{z-4}{3}$$

and the equation of the normal plane is

$$12x - 4y + 3z - 12 = 0$$

To compute the arc length for a space curve determined by equations of form (*) we should solve these equations with respect to two coordinates, for instance, with respect to y and z ; this leads to some expressions of y and z as functions of the coordinate x : $y=y(x)$, $z=z(x)$. After $y(x)$ and $z(x)$ have been determined we can apply the formula for the arc length of a space curve specified parametrically (see Sec. 102), the role of the parameter being played by x : $t=x$. Thus we obtain

$$L = \int_{x_1}^{x_2} \sqrt{1 + y'^2 + z'^2} dx$$

where x_1 and x_2 are the abscissas of the initial and terminal points of the given arc.

Example. Let us compute the length of the arc of the curve specified by the equations

$$x^2 - 3y = 0, \quad 2xy - 9z = 0$$

and joining the point $A(0, 0, 0)$ to the point $B(3, 3, 2)$ (this curve is the line of intersection of the parabolic cylinder $x^2 - 3y = 0$ and the hyperbolic paraboloid $2xy - 9z = 0$).

Expressing y and z from these equations in terms of x we get $y = \frac{x^2}{3}$, $z = \frac{2x^2}{27}$ whence

$$L = \int_0^3 \sqrt{1 + \frac{4x^2}{9} + \frac{4x^4}{81}} dx = \int_0^3 \left(1 + \frac{2x^2}{9}\right) dx = \left(x + \frac{2x^3}{27}\right) \Big|_0^3 = 5$$

§ 4. Extrema of Functions of Two Variables

121. Necessary Conditions for Extremum. The definition of a point of extremum of a function of two variables is analogous to the corresponding definition for a function of one variable (see Sec. 60). Let $z=f(x, y)$ be a function defined in a domain D and let $P_0(x_0, y_0)$ be an interior point of the domain.

Definition. The point $P_0(x_0, y_0)$ is said to be a *point of extremum* (a *point of maximum* or a *point of minimum*) of the function $z = f(x, y)$ if, respectively, $f(x_0, y_0)$ is the greatest or the least value of the function $f(x, y)$ in a neighbourhood of the point $P_0(x_0, y_0)$.

Thus, $f(x, y) < f(x_0, y_0)$ if P_0 is a point of maximum, and $f(x, y) > f(x_0, y_0)$ if P_0 is a point of minimum.

If the strict inequalities are replaced by non-strict ones (i.e. the symbols $<$ and $>$ are replaced by $\leq$ and $\geq$) the point P_0 is a *point of non-strict extremum*.

The value $f(x_0, y_0)$ of the function is called then an *extremal value* (respectively, a *maximum* or a *minimum value*) of the function. We also say in these cases that the function $f(x, y)$ *attains an extremum* (respectively, a *maximum* or a *minimum*) at the point P_0 . The shape of the surfaces representing a function in the vicinity of its points of extremum is shown in Fig. 150.

We shall begin with establishing necessary conditions for a function $z = f(x, y)$ to attain an extremum at a point $P_0(x_0, y_0)$. The function is supposed to be differentiable (extrema of non-differentiable functions will be treated later on).

Necessary Condition for Extremum. If a differentiable function $z = f(x, y)$ attains an extremum at a point $P_0(x_0, y_0)$ its partial derivatives turn into zero at that point:

$$\left(\frac{\partial z}{\partial x}\right)_{\substack{x=x_0 \\ y=y_0}} = 0, \quad \left(\frac{\partial z}{\partial y}\right)_{\substack{x=x_0 \\ y=y_0}} = 0$$

Proof. Let the function $z = f(x, y)$ have an extremum at the point $P_0(x_0, y_0)$. By the definition of an extremum, the function $z = f(x, y)$ regarded as a function of one variable x for the fixed value $y = y_0$ of the other variable attains an extremum at the point $x = x_0$. As is known, for the function $f(x, y_0)$ to have an extremum for $x = x_0$ it is necessary that the derivative of $f(x, y_0)$ with respect to x should vanish at the point $x = x_0$, that is,

$$\left(\frac{\partial z}{\partial x}\right)_{\substack{x=x_0 \\ y=y_0}} = 0$$

Similarly, the function $z = f(x, y)$ regarded as a function of one argument y for the fixed value $x = x_0$ of the variable x has an extremum for $y = y_0$. Hence, $\left(\frac{\partial z}{\partial y}\right)_{\substack{x=x_0 \\ y=y_0}} = 0$, which is what we wanted to prove.

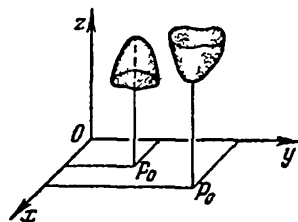


Fig. 150

A point $P_0(x_0, y_0)$ at which both partial derivatives of a function $z=f(x, y)$ turn into zero is referred to as a *stationary point* of the function $f(x, y)$.

The equation of the tangent plane to the surface $z=f(x, y)$ whose general form is

$$z-z_0 = \left(\frac{\partial z}{\partial x}\right)_0 (x-x_0) + \left(\frac{\partial z}{\partial y}\right)_0 (y-y_0)$$

turns into $z=z_0$ for a stationary point $P_0(x_0, y_0)$.

Thus, geometrically, for a differentiable function $z=f(x, y)$ to attain an extremum at a point $P_0(x_0, y_0)$ it is necessary that the tangent plane to the surface (serving as the graph of the function) at the corresponding point should be parallel to the coordinate plane of the independent variables.

To find the stationary points of a given function $z=f(x, y)$ we should equate to zero its both partial derivatives and find the real roots of the resulting equations

$$\frac{\partial z}{\partial x} = 0, \quad \frac{\partial z}{\partial y} = 0 \quad (*)$$

Example. Let us determine the stationary points of the function

$$z = 2x^3 + xy^2 + 5x^2 + y^3$$

In this case the system of equations (*) takes the form

$$6x^2 + y^2 + 10x = 0, \quad 2xy + 2y = 0$$

From the second equation it follows that either $y=0$ or $x=-1$. Substituting, in turn, these values into the first equation we determine the four stationary points

$$M_1(0, 0), \quad M_2\left(-\frac{5}{3}, 0\right), \quad M_3(-1, 2), \quad M_4(-1, -2)$$

To find out which of these points are in fact points of extremum we should apply sufficient conditions for extremum (they will be established in the next section).

It is sometimes possible to examine the character of a stationary point without resorting to sufficient conditions. For instance, if the conditions of the problem directly indicate that the function in question has a maximum or a minimum at some point and if there is only one point satisfying system (*) (i.e. only one pair of values x, y turning equations (*) into identities) this single point is sure to be the sought-for point of extremum.

Finally, note that a continuous function of two variables may have an extremum at a point where it is not differentiable (for instance, such an extremum may correspond to a caspidal point of the surface serving as the graph of the function). As an example, consider the function $z = \sqrt{x^2 + y^2}$. At the origin it obviously

attains a minimum value (equal to zero) but is nondifferentiable for $x=0$, $y=0$; the graph of this function is a circular cone with vertex at the origin and Oz as axis.

Thus, if we include not only differentiable functions but continuous functions as well then

a point of extremum of a function is either a stationary point or a point at which the function is nondifferentiable.

For a function

$$u = f(x_1, x_2, \dots, x_n)$$

of any number n of independent variables the notion of an extremum is defined quite similarly, and the necessary conditions for extremum are derived analogously. Namely:

A differentiable function of n arguments may have an extremum only for those values of $x_1, x_2, \dots, x_n$ for which all its n partial derivatives of the first order are equal to zero:

$$f'_{x_1}(x_1, x_2, \dots, x_n) = 0, \dots, f'_{x_n}(x_1, x_2, \dots, x_n) = 0$$

These equalities form a system of n equations in n unknowns from which the coordinates of the stationary points are found.

122. Sufficient Conditions for Extremum of a Function of Two Variables. As in the case of a function of one variable, the necessary test for an extremum of a function of several variables established in the foregoing section is not sufficient. This means that the fact that the partial derivatives are zero at a given point does not imply that this point is necessarily a point of extremum. For instance, let us consider the function $z = xy$. Its partial derivatives $z'_x = y$ and $z'_y = x$ are equal to zero at the origin where the function has no extremum. For, the function $z = xy$ turns into zero at the origin but attains both positive and negative values in any neighbourhood of the origin (namely, positive values in the first and third quadrants and negative values in the second and fourth quadrants), and therefore the zero value assumed by the function at the origin is neither its maximum nor minimum value.

Sufficient conditions for extremum for a function of several independent variables are essentially more complicated than in the case of a function of one argument. Here we shall state without proof sufficient conditions for extremum for a function of two independent variables.

Let a function $z = f(x, y)$ be continuous together with its partial derivatives of the first and second orders and let $P_0(x_0, y_0)$ be a stationary point of the function, that is,

$$\left(\frac{\partial z}{\partial x}\right)_0 = 0, \quad \left(\frac{\partial z}{\partial y}\right)_0 = 0$$

Let us compute the values of the second derivatives of the function

$f(x, y)$ at the point $P_0(x_0, y_0)$ and denote them, for brevity, A , B and C :

$$A = \left(\frac{\partial^2 z}{\partial x^2} \right)_0, \quad B = \left(\frac{\partial^2 z}{\partial x \partial y} \right)_0, \quad C = \left(\frac{\partial^2 z}{\partial y^2} \right)_0.$$

If $AC - B^2 > 0$ the function $f(x, y)$ has an extremum at the point $P_0(x_0, y_0)$ which is a maximum if $A < 0$ and a minimum if $A > 0$ (the condition $AC - B^2 > 0$ implies that A and C are necessarily of one sign).

If $AC - B^2 < 0$ there is no extremum at the point $P_0(x_0, y_0)$.

If $AC - B^2 = 0$ the properties of the second derivatives do not provide any answer to the question of existence of an extremum, and further investigation is needed.

Examples. (1) We established in Sec. 121 that the function

$$z = 2x^3 + xy^2 + 5x^2 + y^2$$

has the four stationary points

$$M_1(0, 0), \quad M_2\left(-\frac{5}{3}, 0\right), \quad M_3(-1, 2), \quad M_4(-1, -2)$$

The second partial derivatives of the function are

$$\frac{\partial^2 z}{\partial x^2} = 12x + 10, \quad \frac{\partial^2 z}{\partial x \partial y} = 2y, \quad \frac{\partial^2 z}{\partial y^2} = 2x + 2$$

For the point M_1 we have $A = 10$, $B = 0$, $C = 2$. Here $AC - B^2 = 20 > 0$, and hence M_1 is a point of extremum; since A and C are positive it is a point of minimum.

At the point M_2 we have $A = -10$, $B = 0$, $C = -\frac{4}{3}$; this is a point of maximum. Let the reader check that the points M_3 and M_4 are not points of extremum ($AC - B^2 < 0$ for them).

(2) Let us test the function

$$z = x^2 - y^2$$

for extremum. Equating to zero the partial derivatives we receive the equations

$$\frac{\partial z}{\partial x} = 2x = 0, \quad \frac{\partial z}{\partial y} = -2y = 0$$

specifying only one stationary point

at the origin. Here we have $A = 2$, $B = 0$, $C = -2$. Hence $AC - B^2 < 0$, and the point $(0, 0)$ is not a point of extremum. The equation $z = x^2 - y^2$ represents a hyperbolic paraboloid; Fig. 151 clearly indicates that the stationary point $(0, 0)$ is not a point of extremum.

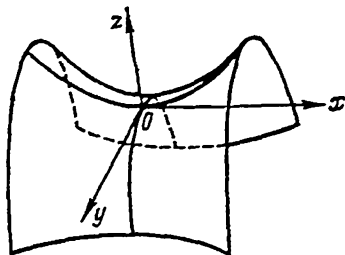


Fig. 151

123. Determining the Greatest and the Least Values of Functions. Suppose that it is required to determine the greatest and the least values of a function $z=f(x, y)$ in a closed domain D . If one of these values (or both) is attained inside the domain it must be of course an extremal value. But it may turn out that the greatest or the least value of the function (or both) is attained at a point belonging to the boundary of the domain. Therefore, *to find the greatest (the least) value of a continuous function $z=f(x, y)$ in a bounded closed domain it is necessary to determine all the maxima (minima) of the function attained inside the domain and also its greatest (least) value on the boundary of the domain. Then the greatest (the least) of these numbers is the sought-for greatest (least) value of the function.*

Example. Let us determine the point in the xy -plane for which the sum of the squares of its distances from the three given points $P_1(0, 0)$, $P_2(1, 0)$ and $P_3(0, 1)$ has the least possible value and also the point within the triangle with vertices at P_1 , P_2 and P_3 for which the sum of the squares of its distances to the vertices has the greatest value.

Take an arbitrary point $P(x, y)$ in the plane. The sum of the squares of its distances from the given points P_1, P_2, P_3 (we denote this sum by z) is expressed by the formula

$$z = x^2 + y^2 + (x-1)^2 + y^2 + x^2 + (y-1)^2$$

i.e.

$$z = 3x^2 + 3y^2 - 2x - 2y + 2$$

The first problem reduces to finding the least value of this function throughout the plane and the second problem to determining its greatest value on condition that the point $P(x, y)$ belongs to the closed domain $\bar{D}$ bounded by the triangle $P_1P_2P_3$.

Let us find the extrema of the function $z = 3x^2 + 3y^2 - 2x - 2y + 2$. From the equations

$$\frac{\partial z}{\partial x} = 6x - 2 = 0 \quad \text{and} \quad \frac{\partial z}{\partial y} = 6y - 2 = 0$$

we obtain

$$x = \frac{1}{3}, \quad y = \frac{1}{3}$$

Hence, there is only one stationary point $P\left(\frac{1}{3}, \frac{1}{3}\right)$. The function z considered throughout the plane has no greatest value: since it is clear that, given an arbitrary number, there are points in the plane for which the sum of their distances to the origin exceeds this number. On the other hand, this sum obviously possesses the least value, and therefore the stationary point P is the only one at which the function attains its least value (equal

to $\frac{1}{3}$). By the way, the point $P\left(\frac{1}{3}, \frac{1}{3}\right)$ is the centre of gravity of the triangle with vertices at P_1, P_2, P_3 .

Now we proceed to the second problem. The given function having no maximum, its greatest value in the domain D (the triangle $P_1P_2P_3$) is the greatest of the values attained on the boundary of D , that is, on the sides of the triangle.

On the side P_1P_2 , we have $y=0$, and hence,

$$z = 3x^2 - 2x + 2$$

This function attains its greatest value (equal to 3) within the interval $0 \leq x \leq 1$ for $x=1$, that is, at the point P_2 .

On the side P_1P_3 , we have $x=0$, and thus

$$z = 3y^2 - 2y + 2$$

The greatest value of this function in the interval $0 \leq y \leq 1$ is attained for $y=1$, i.e. at the point P_3 , and is also equal to 3.

Finally, for the side P_2P_3 , we have $x+y=1$; consequently,

$$z = 3x^2 + 3(1-x)^2 - 2x - 2(1-x) + 2 = 6x^2 - 6x + 3$$

The latter function achieves its maximum value (equal to 3) in the interval $0 \leq x \leq 1$ for $x=0$ and $x=1$; this means that the greatest value on the side P_2P_3 is attained at the same points P_2 and P_3 . Thus, for the second problem we receive two points: P_2 and P_3 . Among all the points of the triangle these two points have the greatest value of the sum of the squares of their distances to the vertices P_1, P_2, P_3 .

124*. Conditional Extremum. In problems involving the determination of extrema of functions of several variables we often encounter the so-called conditional extrema. This notion will be demonstrated by the example of a function of two arguments.

Let there be given a function $z=f(x, y)$ and a curve L in the xy -plane. The problem is to find a point $P(x, y)$ on the curve L such that the value $f(x, y)$ of the function at that point is the greatest or the least compared with the values of the function at the points of the line L lying in the vicinity of the point P . A point P of this kind is termed a *point of conditional extremum of the function $f(x, y)$ on the line L* . The distinction from an ordinary point of extremum (sometimes referred to as a *point of unconditional extremum*) is that the value of the function at a point of conditional extremum is compared not with the values at all the points of its sufficiently small neighbourhood but only at those belonging to the curve L .

It is obvious that a point of unconditional extremum is always a point of conditional extremum for any curve passing through this point. Of course, the converse is not true: a point of condi-

tional extremum is not necessarily a point of ordinary extremum. What has been said can be illustrated by the following simple example. The graph of the function $z = \sqrt{1-x^2-y^2}$ is the upper hemisphere of radius 1 with centre at the origin (see Fig. 152). The function has a maximum at the origin, the corresponding point being the highest point M of the hemisphere. If L is the straight line passing through the points $A(1, 0)$ and $B(0, 1)$ (its equation is $x+y-1=0$) it is clear, geometrically, that for the points of this line the greatest value of the function is attained at the point $P\left(\frac{1}{2}, \frac{1}{2}\right)$ bisecting the line segment AB . It is this point at which the conditional extremum (maximum) of the function $z = \sqrt{1-x^2-y^2}$ is achieved on the given line, the corresponding point of the hemisphere being M_1 . At the same time it is apparent that there is no ordinary extremum at that point.

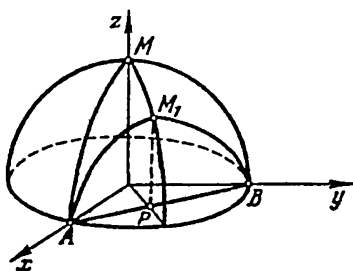


Fig. 152

Note that in the problem of determining the greatest and the least values of a function in a closed domain considered in Sec. 123 we dealt with extremal values of the function on the boundary of the domain, that is, on a certain curve, and thus we were concerned with a problem of conditional extremum.

Now we proceed to practical determination of the points of conditional extremum of a given function $z = f(x, y)$ on a curve specified by an equation of the form $\varphi(x, y) = 0$. Such an equation is called a *constraint equation* or, simply, a *constraint* (also a *coupling equation* or a *subsidiary condition*). If this equation admits of the expressions of y as an explicit function of x , i.e. $y = \psi(x)$, we can substitute $\psi(x)$ for y into the function $z = f(x, y)$ to obtain the function of one variable

$$z = f[x, \psi(x)] = \Phi(x)$$

On finding the values of x for which this function attains extrema and determining the corresponding values of y from the equation $\varphi(x, y) = 0$ we obtain the desired points of conditional extremum.

For instance, in the above example of the conditional extremum of the function $z = \sqrt{1-x^2-y^2}$ with constraint $x+y-1=0$ we have $y = 1-x$, whence

$$z = \sqrt{1-x^2-(1-x)^2} = \sqrt{2x-2x^2}$$

It is readily checked that z has a maximum for $x=0.5$; the subsidiary condition $x+y-1=0$ then yields the value $y=0.5$, and

we thus obtain the same point P which was earlier found geometrically.

A conditional extremum problem can be solved particularly simply when the subsidiary condition is representable by parametric equations $x=x(t)$, $y=y(t)$. For, on substituting the expressions of x and y in terms of t into the given function we immediately arrive at a problem of determining an extremum of a function of one variable.

If the subsidiary condition is expressed by a complicated equation and if it is impossible to express explicitly one variable in terms of the other or to pass to parametric equations the problem becomes more difficult. To analyse this case let us consider, as before, the variable y (entering into the expression $z=f(x, y)$) as a function of x specified implicitly by the equation $\varphi(x, y)=0$. The total derivative of the function $z=f(x, y)$ with respect to x (see Sec. 116, I) is then written as

$$\frac{dz}{dx} = f'_x(x, y) + f'_y(x, y) \frac{dy}{dx} = f'_x(x, y) - \frac{\varphi'_x(x, y)}{\varphi'_y(x, y)} f'_y(x, y)$$

where y' has been found according to the differentiation rule for an implicit function (Sec. 118). The total derivative thus found must turn into zero at the points of conditional extremum, which provides an equation connecting x and y . These variables must also satisfy the subsidiary equation, and hence we arrive at the system of two equations in two unknowns:

$$f'_x - \frac{\varphi'_x}{\varphi'_y} f'_y = 0, \quad \varphi(x, y) = 0$$

Now let us bring this system to a more convenient form by introducing a new auxiliary variable λ and writing the first equation as the proportion

$$\frac{f'_x}{\varphi'_x} = \frac{f'_y}{\varphi'_y} = -\lambda \quad (*)$$

(the minus sign in λ is chosen for the sake of convenience). These relations imply the equations

$$f'_x(x, y) + \lambda \varphi'_x(x, y) = 0, \quad f'_y(x, y) + \lambda \varphi'_y(x, y) = 0 \quad (**)$$

which, together with the equation of constraint $\varphi(x, y)=0$, form a system of equations in the three unknowns x , y and λ .

To memorize equations. (**) it is convenient to state them as the following rule:

To determine the points at which a conditional extremum can be attained by a function $z=f(x, y)$ with constraint $\varphi(x, y)=0$ one should form the auxiliary function

$$\Phi(x, y) = f(x, y) + \lambda \varphi(x, y)$$

where λ is an arbitrary constant parameter and solve the equations for the points of unconditional extremum for this auxiliary function.

As a rule, this system of equations only provides *necessary conditions*, that is, not every pair of values of x and y satisfying this system is necessarily a point of conditional extremum. In our course we shall not discuss *sufficient conditions* for the points of conditional extremum; in a concrete problem the given conditions often make it possible to find out whether the point determined from the above equations is an extremal point without resorting to sufficient conditions. The parameter λ introduced above is called *Lagrange's multiplier*, and the method itself is known as the *method of Lagrange's multipliers*.

Lagrange's method admits of a simple geometrical interpretation which we are going to discuss now. Suppose that the level lines of the function $z = f(x, y)$ (see Sec. 107) in question are as shown in Fig. 153, the curve L being the one for which the conditional extremum is sought.

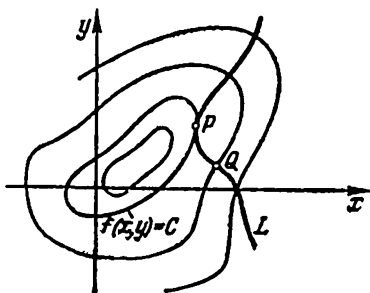


Fig. 153

If the line L intersects a level line at a point Q this point cannot be a point of conditional extremum since the function $f(x, y)$ assumes the values greater than that at the point Q on one side of the level line and smaller values on the other side. But if the line L has a common point P with a level line and does not intersect the latter it lies entirely on one side of that level line in the vicinity of the point, and then this point is a point of conditional extremum. At such a point the curve L and the level line in question $f(x, y) = C$ are tangent to each other (the lines are supposed to be smooth), and the slopes of their tangents must coincide. From the coupling equation $\varphi(x, y) = 0$ we get $y' = -\varphi'_x/\varphi'_y$, while the equation of the level line implies $y' = -f'_x/f'_y$. On equating the derivatives and making simple transformations we arrive at equations (*).

This argument may prove inapplicable if the level line in question is such that $f'_x = 0$ and $f'_y = 0$ at all its points. As an example, let the reader consider the function $z = 4 - x^2$ and the level line $x = 0$ corresponding to the value $z = 4$.

Example. Let us determine the greatest value of the function $z = xy$ on condition that x and y are positive and satisfy the subsidiary equation $\frac{x^2}{8} + \frac{y^2}{2} - 1 = 0$. Here we have a simple constraint (specifying an ellipse), and therefore the problem can be

reduced directly to the determination of an extremum of a function of one variable if y is expressed in terms of x or if the ellipse is represented parametrically. Let the reader carry out these calculations; to demonstrate the method of Lagrange's multipliers we shall apply to this problem the general techniques described above.

Let us form the auxiliary function:

$$\Phi(x, y) = xy + \lambda \left(\frac{x^2}{8} + \frac{y^2}{2} \right)$$

(The constant term on the left-hand side of the subsidiary equation gives zero when differentiated, and therefore we have simply discarded it.) On equating to zero the partial derivatives with respect to x and y we obtain

$$\frac{\partial \Phi}{\partial x} = y + \frac{\lambda x}{4} = 0, \quad \frac{\partial \Phi}{\partial y} = x + \lambda y = 0$$

The elimination of λ now leads to the equation $4y^2 - x^2 = 0$ which should be solved simultaneously with the coupling equation. This solution is $x=2$ and $y=1$ (the conditions of the problem require that x and y should be positive).

For the values of x and y in question the function $z=xy$ is positive and turns into zero at the points $(2\sqrt{2}, 0)$ and $(0, \sqrt{2})$ where the ellipse intersects the coordinate axes. Therefore the single stationary point $P(2, 1)$ of the auxiliary function is the point of conditional extremum of the given function. Namely, at this point the function attains its conditional maximum: $z_{\max} = 2$. It can readily be verified that the ellipse touches at the point P the level line of the function $z=xy$ passing through that point (this level line is the hyperbola $xy=2$).

There is a wider variety of conditional extremum problems for functions of three variables. Let there be given a function $u=f(x, y, z)$ of three variables x, y and z connected by one equation $\varphi(x, y, z)=0$ which is an equation of a surface. Then a point of conditional extremum is a point of the surface $\varphi(x, y, z)=0$ at which the given function attains the greatest or the least value compared with its values at all the points on the surface in the vicinity of that point.

We can also pose the extremum problem for the function $f(x, y, z)$ on condition that the variables x and y are connected by two constraints: $\varphi_1(x, y, z)=0$ and $\varphi_2(x, y, z)=0$.¹ These subsidiary conditions specify a curve in space, and hence the problem reduces to determining a point of the curve at which

¹ Let the reader explain why the problem may lose its sense for a greater number of constraints.

the function assumes its extremal value, only the values of the function at the points of that curve being compared.

In the case of two constraints the method of Lagrange's multipliers is applied as follows. We construct the auxiliary function

$$\Phi(x, y, z) = f(x, y, z) + \lambda_1 \varphi_1(x, y, z) + \lambda_2 \varphi_2(x, y, z)$$

with two additional unknown parameters (Lagrange's multipliers) λ_1 and λ_2 and form the system of equations for determining the points of unconditional extremum of this new function:

$$\frac{\partial \Phi}{\partial x} = \frac{\partial f}{\partial x} + \lambda_1 \frac{\partial \varphi_1}{\partial x} + \lambda_2 \frac{\partial \varphi_2}{\partial x} = 0$$

$$\frac{\partial \Phi}{\partial y} = \frac{\partial f}{\partial y} + \lambda_1 \frac{\partial \varphi_1}{\partial y} + \lambda_2 \frac{\partial \varphi_2}{\partial y} = 0$$

$$\frac{\partial \Phi}{\partial z} = \frac{\partial f}{\partial z} + \lambda_1 \frac{\partial \varphi_1}{\partial z} + \lambda_2 \frac{\partial \varphi_2}{\partial z} = 0$$

Adding to these equations the two subsidiary conditions we arrive at a system of five simultaneous equations in the five unknowns $x, y, z, \lambda_1, \lambda_2$. The sought-for points of conditional extremum must be among those whose coordinates x, y, z , are solutions of this system.

In the most general case the problem is posed as follows: given a function $u = f(x_1, x_2, \dots, x_n)$ of n variables, it is required to find its extrema on condition that the variables are subjected to m ($m < n$) subsidiary conditions

$$\varphi_1(x_1, x_2, \dots, x_n) = 0$$

$$\varphi_2(x_1, x_2, \dots, x_n) = 0, \dots, \varphi_m(x_1, x_2, \dots, x_n) = 0$$

In this case the auxiliary function of n variables involves m additional unknown parameters (Lagrange's multipliers):

$$\Phi = f + \lambda_1 \varphi_1 + \lambda_2 \varphi_2 + \dots + \lambda_m \varphi_m$$

To find the points of extrema of this function we form the corresponding system of $m+n$ equations from which the possible values of the coordinates $x_1, x_2, \dots, x_n$ of the points of conditional extremum are found.

Examples. (1) Among all the rectangular parallelepipeds with given total surface area S it is necessary to choose the one of the greatest volume.

Let the sides of the parallelepiped be denoted x, y and z . The problem reduces to determining the greatest value of the function $V = xyz$ on condition that $xy + yz + zx = \frac{S}{2}$.

The auxiliary function is $\Phi(x, y, z) = xyz + \lambda(xy + yz + zx)$. The equations for determining the points of extremum are of the form

$$yz + \lambda(y + z) = 0, \quad xz + \lambda(x + z) = 0, \quad xy + \lambda(x + y) = 0$$

On subtracting these equations from one another we receive

$$\begin{aligned}(z + \lambda)(y - x) &= 0, & (x + \lambda)(z - y) &= 0, \\ (y + \lambda)(x - z) &= 0\end{aligned}$$

It follows that $x = y = z$, that is, the sought-for parallelepiped is a cube.

The dimensions of this cube are found from the constraints:

$$x = y = z = \sqrt{\frac{S}{6}} \quad \text{and} \quad V = \frac{S \sqrt{S}}{6 \sqrt{6}}$$

(2) Let us determine the greatest distance from the origin to the points of the line of intersection of the paraboloid of revolution $z = x^2 + y^2$ with the plane $x + 2y - z = 0$. (The least distance is equal to zero since the line of intersection passes through the origin.)

The distance from a point $P(x, y, z)$ to the origin is equal to $r = \sqrt{x^2 + y^2 + z^2}$. The greatest value of the root and that of the radicand are attained at the same point; therefore we shall find the conditional maximum of the function $u = x^2 + y^2 + z^2$ for two subsidiary conditions

$$\varphi_1(x, y, z) \equiv x^2 + y^2 - z = 0 \quad \text{and} \quad \varphi_2(x, y, z) \equiv x + 2y - z = 0$$

On forming the auxiliary function

$$\Phi = x^2 + y^2 + z^2 + \lambda_1(x^2 + y^2 - z) + \lambda_2(x + 2y - z)$$

and equating to zero its partial derivatives we obtain

$$2x + 2\lambda_1 x + \lambda_2 = 0, \quad 2y + 2\lambda_1 y + 2\lambda_2 = 0, \quad 2z - \lambda_1 - \lambda_2 = 0$$

Let us express the coordinates x , y and z from these relations in terms of Lagrange's multipliers:

$$x = -\frac{\lambda_2}{2(1 + \lambda_1)}, \quad y = -\frac{\lambda_2}{1 + \lambda_1}, \quad z = \frac{\lambda_1 + \lambda_2}{2}$$

Substituting these expressions into the subsidiary conditions we receive

$$\frac{5}{4} \left(\frac{\lambda_2}{1 + \lambda_1} \right)^2 = \frac{\lambda_1 + \lambda_2}{2}, \quad -\frac{5}{2} \left(\frac{\lambda_2}{1 + \lambda_1} \right) = \frac{\lambda_1 + \lambda_2}{2}$$

Finally, on equating the left-hand sides of these equations we arrive at the quadratic equation with respect to the ratio $\frac{\lambda_2}{1 + \lambda_1}$. Next we find the two roots of this quadratic equation and then determine the values of λ_1 and λ_2 . One of the roots is $\frac{\lambda_2}{1 + \lambda_1} = 0$; it leads to the values $\lambda_1 = 0$ and $\lambda_2 = 0$. The point corresponding to them is the origin at which a minimum is achieved. The other

root is $\frac{\lambda_2}{1+\lambda_1} = -2$. The values of x , y and z corresponding to it are found directly without computing λ_1 and λ_2 : $x=1$, $y=2$ and $z=5$. It is clear that the point $(1, 2, 5)$ is at the greatest distance from the origin, and $r_{\max} = \sqrt{30}$.

§ 5. Scalar Field

125. The Notion of a Scalar Field. Level Surfaces. Suppose that at each point P of a domain D a certain value of a scalar physical magnitude u is specified; a magnitude of this kind is completely determined by its numerical values. Examples of such magnitudes are the temperature at the points of a heated body which may vary from point to point, the density of the distribution of an electric charge in an electrified body, the potential of an electric field, etc. In such a case we say that there is a *scalar point function* defined in the domain D and write $u=u(P)$.

The domain D in which the function $u(P)$ is defined may coincide with the whole space or be its part.

Definition. If there is a scalar point function $u(P)$ defined in a domain D it is called a *scalar field in D* .

We shall suppose that the field in question is *stationary*, that is, the magnitude $u(P)$ is independent of time t . By the way, later on (see Sec. 157) we shall also encounter *non-stationary* fields, and then the magnitude u will depend not only on the point P but also on time t .

If the physical magnitude in question is a *vector quantity* the corresponding field is spoken of as a *vector field*; such are, for example, a field of force, an electric field (i.e. a field of electric intensity), a magnetic field and the like. We shall treat vector fields in § 3, Chapter IX.

If a scalar field is related to a coordinate system $Oxyz$ the specification of the variable point P is equivalent to the specification of its coordinates x , y , z , and the function $u(P)$ can then be written as an ordinary function of three variables: $u=u(x, y, z)$. We have thus arrived at a physical interpretation of a function of three variables.

Definition. A *level surface* of a scalar field is the locus of points at which the corresponding function u assumes a given constant value, that is,

$$u(x, y, z) = C$$

In physics, when dealing with a field of a potential, we consider the corresponding level surfaces, that is, surfaces of constant potential, referred to as *equipotential surfaces*.

The equation of the level surface passing through a given point $M_0(x_0, y_0, z_0)$ is written

$$u(x, y, z) = u(x_0, y_0, z_0)$$

In particular, if the scalar field in question is *plane*, that is, if we deal with a distribution of a physical magnitude in a plane region, the function u depends on two variables, for instance, on x and y . Then we speak of *level lines* of the field which are the level lines of the function $u(x, y)$ (see Sec. 108, I):

$$u(x, y) = C$$

126. Directional Derivative. An important characteristic of a scalar field is the rate of its change in the given direction. Consider a scalar field, i.e. a function $u(x, y, z)$. Let us take a point $P(x, y, z)$ and a ray λ issued from it. The direction of the ray can be specified by the angles α, β, γ it forms with the positive directions of the coordinate axes Ox, Oy, Oz (Fig. 154). If e_λ is unit vector along the ray λ its projections are the direction cosines:

$$e_\lambda = \{\cos \alpha, \cos \beta, \cos \gamma\}$$

Let $P_1(x_1, y_1, z_1)$ be a point on the ray λ , the distance PP_1 being denoted ρ . The projections of the vector $\overline{PP_1}$ on the coordinate axes are, on the one hand, equal to $\rho \cos \alpha, \rho \cos \beta$ and $\rho \cos \gamma$, and, on the other hand, to the differences $x_1 - x, y_1 - y$ and $z_1 - z$. Consequently,

$$x_1 = x + \rho \cos \alpha, \quad y_1 = y + \rho \cos \beta, \quad z_1 = z + \rho \cos \gamma$$

Now let us consider the increment of the function u gained as we pass from the point P to the point P_1 :

$$u(P_1) - u(P) = u(x + \rho \cos \alpha, y + \rho \cos \beta, z + \rho \cos \gamma) - u(x, y, z)$$

If the point P_1 moves in the ray λ this results solely in a change of the magnitude ρ in the expression of the difference $u(P_1) - u(P)$. Let us form the ratio

$$\frac{u(P_1) - u(P)}{\rho}$$

and pass to the limit as $\rho \rightarrow 0$. This leads to the notion of the (directional) derivative along a given direction:

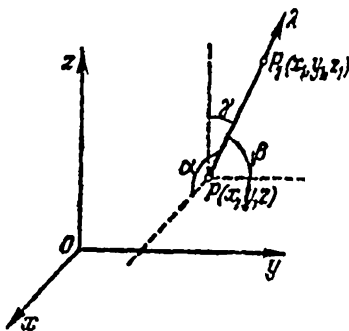


Fig. 154

Definition. The (*directional*) *derivative* of the function $u(x, y, z)$ in the direction of the ray λ at the point P is the limit

$$\lim_{\rho \rightarrow 0} \frac{u(P_1) - u(P)}{\rho} = \lim_{\rho \rightarrow 0} \frac{u(x + \rho \cos \alpha, y + \rho \cos \beta, z + \rho \cos \gamma) - u(x, y, z)}{\rho} \quad (*)$$

(provided it exists).

We shall denote this limit by $\frac{\partial u}{\partial \lambda}$ or $u'_\lambda(x, y, z)$. Its magnitude depends both on the choice of the point $P(x, y, z)$ and on the direction of the ray λ , i.e. on α, β and γ . If the position of the point P is fixed the value of the directional derivative u'_λ only depends on the direction of the ray.

The definition of the directional derivative indicates that if the direction of the ray λ coincides with the positive direction of the x -axis, that is, if $\alpha = 0$ and $\beta = \gamma = \frac{\pi}{2}$, limit $(*)$ is simply equal to the partial derivative of the function $u(x, y, z)$ with respect to x :

$$\lim_{\rho \rightarrow 0} \frac{u(x + \rho, y, z) - u(x, y, z)}{\rho} = \frac{\partial u}{\partial x}$$

An analogous situation occurs when λ goes along Oy or Oz .

As is known, the partial derivatives u'_x, u'_y and u'_z characterize the rate of change of the function u in the directions of the coordinate axes (see Sec. 110). Similarly, the directional derivative u'_λ is equal to the *rate of change* of the function $u(x, y, z)$ at the point P in the direction of the ray λ . The *absolute value* of the derivative u'_λ in the direction of the ray λ determines the magnitude of the rate of change in this direction, and the *sign of the derivative* indicates the character of the variation of u (i.e. its increase or decrease).

The computation of the directional derivative is based on the following theorem:

Theorem. If $u(x, y, z)$ is a differentiable function its directional derivative exists for the ray λ going in any direction and is expressed by the formula

$$\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma \quad (**)$$

where $\cos \alpha, \cos \beta$ and $\cos \gamma$ are the direction cosines of the ray λ .

Proof. The function $u(x, y, z)$ being differentiable, its total increment (see Sec. 111, I) can be written as

$$\begin{aligned} \Delta u &= u(x + \Delta x, y + \Delta y, z + \Delta z) - u(x, y, z) = \\ &= \frac{\partial u}{\partial x} \Delta x + \frac{\partial u}{\partial y} \Delta y + \frac{\partial u}{\partial z} \Delta z + o(\rho) \end{aligned}$$

where $\rho = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}$. Putting $\Delta x = \rho \cos \alpha$, $\Delta y = \rho \cos \beta$ and $\Delta z = \rho \cos \gamma$ we can represent the difference $u(P_1) - u(P)$ in the form

$$u(P_1) - u(P) = \frac{\partial u}{\partial x} \rho \cos \alpha + \frac{\partial u}{\partial y} \rho \cos \beta + \frac{\partial u}{\partial z} \rho \cos \gamma + o(\rho)$$

Hence, the ratio $\frac{u(P_1) - u(P)}{\rho}$ is expressed as

$$\frac{u(P_1) - u(P)}{\rho} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma + \frac{o(\rho)}{\rho}$$

The values of the partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial z}$ at the point P and also α , β and γ being independent of ρ , the passage to the limit as $\rho \rightarrow 0$ leads to

$$\frac{\partial u}{\partial \lambda} = \lim_{\rho \rightarrow 0} \frac{u(P_1) - u(P)}{\rho} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma$$

which is what we intended to prove.

Formula (**) implies directly the above assertion that if the direction of the ray λ coincides with the positive direction of one of the coordinate axes the derivative in this direction is equal to the corresponding partial derivative; for instance, if $\alpha = 0$ and $\beta = \gamma = \frac{\pi}{2}$ then $\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial x}$.

Formula (**) also shows that the derivative in the direction λ' opposite to λ is equal to the minus derivative along the ray λ . In fact, when the direction is changed to the opposite the angles α , β and γ gain the increments equal to π , and therefore

$$\frac{\partial u}{\partial \lambda'} = \frac{\partial u}{\partial x} \cos(\alpha + \pi) + \frac{\partial u}{\partial y} \cos(\beta + \pi) + \frac{\partial u}{\partial z} \cos(\gamma + \pi) = -\frac{\partial u}{\partial \lambda}$$

Thus, if the direction is reversed the absolute value of the rate of change of the function u remains the same, and only the character of variation of the function is changed: if, for example, the function increases in the direction λ it decreases in the opposite direction λ' and vice versa.

Example. Consider the function $u = xyz$. Let us determine its derivative at the point $P(5, 1, 2)$ in the direction from this point to the point $Q(7, -1, 3)$.

We first find the partial derivatives of the function $u = xyz$:

$$\frac{\partial u}{\partial x} = yz, \quad \frac{\partial u}{\partial y} = xz, \quad \frac{\partial u}{\partial z} = xy$$

Now, it is necessary to compute their values at the point P :

$$\left(\frac{\partial u}{\partial x}\right)_P = 2, \quad \left(\frac{\partial u}{\partial y}\right)_P = 10, \quad \left(\frac{\partial u}{\partial z}\right)_P = 5$$

The projections of the vector $\overline{PQ}$ being equal to 2, -2 and 1, its direction cosines are

$$\cos \alpha = \frac{2}{\sqrt{2^2 + (-2)^2 + 1^2}} = \frac{2}{3}, \quad \cos \beta = \frac{-2}{3}, \quad \cos \gamma = \frac{1}{3}$$

Consequently,

$$\frac{\partial u}{\partial \lambda} = 2 \cdot \frac{2}{3} + 10 \cdot \left(-\frac{2}{3}\right) + 5 \cdot \frac{1}{3} = -\frac{11}{3}$$

The minus sign indicates that the function u decreases in the given direction.

If we deal with a *plane* field the direction of a given ray λ is completely determined by its angle of inclination α to the axis of abscissas. The formula for the directional derivative of a plane field can be obtained from the general formula (**) by putting $\gamma = \frac{\pi}{2}$ and $\beta = \frac{\pi}{2} - \alpha$, which yields

$$\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \sin \alpha$$

If $\alpha = 0$ then $\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial x}$, and if $\alpha = \frac{\pi}{2}$ then $\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial y}$.

127. Gradient. Let us examine more closely the formula for the directional derivative:

$$\frac{\partial u}{\partial \lambda} = \frac{\partial u}{\partial x} \cos \alpha + \frac{\partial u}{\partial y} \cos \beta + \frac{\partial u}{\partial z} \cos \gamma$$

As was mentioned, the second factors in each summand are the projections of the unit vector e_λ of the ray λ :

$$e_\lambda = \{\cos \alpha, \cos \beta, \cos \gamma\}$$

Now let us take the vector whose projections on the coordinate axes are equal to the values of the partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial z}$ at a chosen point $P(x, y, z)$. This vector is called the *gradient* of the function $u(x, y, z)$ at that point; it is denoted by the symbol $\text{grad } u$ or ∇u .

Definition. The *gradient* of the function $u(x, y, z)$ is the vector whose projections are the values of the partial derivatives of the function, that is,

$$\text{grad } u = \frac{\partial u}{\partial x} \mathbf{i} + \frac{\partial u}{\partial y} \mathbf{j} + \frac{\partial u}{\partial z} \mathbf{k}$$

It should be stressed that the projections of the gradient depend on the choice of the point $P(x, y, z)$ and may vary when the coordinates of this point change. Thus, at each point of the domain

of definition of the scalar field specified by the function $u(x, y, z)$ there is a definite vector equal to the gradient of that function. Note that the gradient of the linear function $u = ax + by + cz + d$ is a constant vector: $\text{grad } u = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$.

Using the notion of the gradient we can rewrite the formula for the directional derivative in the form

$$\frac{\partial u}{\partial \lambda} = \text{grad } u \cdot \mathbf{e}_\lambda \quad (*)$$

Consequently,

the derivative of a function in a given direction is equal to the scalar product of the gradient of the function by the unit vector in that direction.

The scalar product of two vectors being equal to the modulus of one vector multiplied by the projection of the other vector on the former, we can also state that

the derivative of a function in a given direction is equal to the projection of the gradient of the function on the ray along which the differentiation is carried out, that is

$$\frac{\partial u}{\partial \lambda} = |\text{grad } u| \cos \varphi$$

where φ is the angle between the vector $\text{grad } u$ and the ray λ . (Fig. 155).

It follows immediately that the directional derivative attains its greatest value for $\cos \varphi = 1$,

i.e. for $\varphi = 0$, this greatest value being equal to $|\text{grad } u|$.

Thus, $|\text{grad } u|$ is the greatest possible value of the derivative u_λ at the given point P , and the direction of the vector $\text{grad } u$ coincides with the direction of the ray issued from the point P along which the rate of change of the function is the greatest, that is, the direction of the gradient is that of the fastest increase of the function. It is obvious that the opposite direction is the one along which the function has the fastest decrease.

Now we proceed to prove a theorem establishing the relationship between the gradient of a function and the level surfaces of the corresponding scalar field.

Theorem. The gradient of a function $u(x, y, z)$ is at each point in the direction of the normal to the level surface, of the corresponding scalar field, passing through that point (see Fig. 156).

Proof. Let us choose an arbitrary point $P_0(x_0, y_0, z_0)$. The equation of the level surface passing through the point P_0 is written as $u(x, y, z) = u_0$ where $u_0 = u(x_0, y_0, z_0)$.

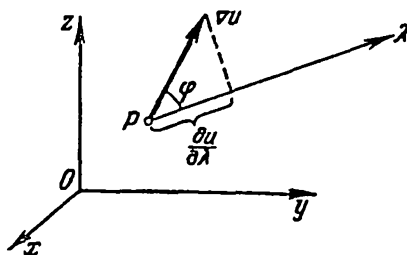


Fig. 155

Let us form the equation of the normal to this surface at the point P_0 (see Sec. 119):

$$\frac{x-x_0}{\left(\frac{\partial u}{\partial x}\right)_0} = \frac{y-y_0}{\left(\frac{\partial u}{\partial y}\right)_0} = \frac{z-z_0}{\left(\frac{\partial u}{\partial z}\right)_0}$$

This equation shows directly that the gradient of the function $u(x, y, z)$ at the point P_0 (its projections are $\left(\frac{\partial u}{\partial x}\right)_0$, $\left(\frac{\partial u}{\partial y}\right)_0$ and $\left(\frac{\partial u}{\partial z}\right)_0$) is a direction vector of the normal, which is what we set out to prove.

Thus, the gradient is at each point perpendicular to the tangent plane to the level surface passing through the given point¹, i.e. its projection on that plane is equal to zero. Thus:

the derivative in any direction tangent to the level surface passing through the given point is equal to zero.

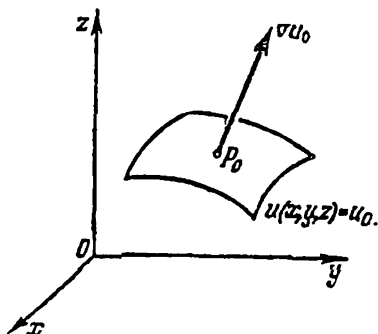


Fig. 156

We shall enumerate some properties of the gradient which often facilitate its computation.

(1) $\text{grad } (u_1 + u_2) = \text{grad } u_1 + \text{grad } u_2.$

(2) $\text{grad } Cu_1 = C \text{ grad } u_1$ where C is a constant.

Let the reader establish these properties.

(3) $\text{grad } u_1 u_2 = u_2 \text{ grad } u_1 + u_1 \text{ grad } u_2.$ Indeed, we have

$$\begin{aligned} \text{grad } u_1 u_2 &= \frac{\partial (u_1 u_2)}{\partial x} \mathbf{i} + \frac{\partial (u_1 u_2)}{\partial y} \mathbf{j} + \frac{\partial (u_1 u_2)}{\partial z} \mathbf{k} = \\ &= u_2 \left(\frac{\partial u_1}{\partial x} \mathbf{i} + \frac{\partial u_1}{\partial y} \mathbf{j} + \frac{\partial u_1}{\partial z} \mathbf{k} \right) + u_1 \left(\frac{\partial u_2}{\partial x} \mathbf{i} + \frac{\partial u_2}{\partial y} \mathbf{j} + \frac{\partial u_2}{\partial z} \mathbf{k} \right) = \\ &= u_2 \text{ grad } u_1 + u_1 \text{ grad } u_2. \end{aligned}$$

(4) $\text{grad } f(u) = f'(u) \text{ grad } u.$ For we have

$$\begin{aligned} \text{grad } f(u) &= \frac{\partial}{\partial x} [f(u)] \mathbf{i} + \frac{\partial}{\partial y} [f(u)] \mathbf{j} + \frac{\partial}{\partial z} [f(u)] \mathbf{k} = \\ &= f'(u) \frac{\partial u}{\partial x} \mathbf{i} + f'(u) \frac{\partial u}{\partial y} \mathbf{j} + f'(u) \frac{\partial u}{\partial z} \mathbf{k} = f'(u) \text{ grad } u \end{aligned}$$

These properties of the gradient indicate that it can be computed by analogy with the well-known differentiation rules for functions.

¹ If there is a point at which $\text{grad } u = 0$ it is a singular point of the level surface (see Sec. 119); as was said, such points are not treated in our course.

Examples. (1) Let us take the function $r = \sqrt{x^2 + y^2 + z^2}$ (expressing the distance from the point (x, y, z) to the origin). We have

$$\text{grad } r \frac{\partial r}{\partial x} \mathbf{i} + \frac{\partial r}{\partial y} \mathbf{j} + \frac{\partial r}{\partial z} \mathbf{k} = \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{r} = \frac{\mathbf{r}}{r}$$

and hence, the vector $\text{grad } r$ is in the direction of the radius vector $\mathbf{r}$, its modulus being equal to unity.

(2) Consider the scalar field specified by the function $\frac{q}{r}$ where r is the same as in Example 1 and $q = \text{const.}$ According to Property (4), we write

$$\text{grad } \frac{q}{r} = -\frac{q}{r^2} \text{grad } r = -\frac{q}{r^2} \frac{\mathbf{r}}{r}$$

For a plane field $u = u(x, y)$ its gradient

$$\text{grad } u = \frac{\partial u}{\partial x} \mathbf{i} + \frac{\partial u}{\partial y} \mathbf{j}$$

lies in the xy -plane and is perpendicular to the level line passing through the given point (x, y) .

If a sufficiently dense system of contour lines of a given plane field (see Fig. 157) has been constructed, it is possible to obtain approximations to the modulus and to the direction of the gradient. Namely, the direction of the gradient is always perpendicular to the corresponding level line, and, for sufficiently small values of h , the derivative in this direction is approximately given by the formula

$$\frac{\partial u}{\partial \lambda} \approx \frac{u(P) - u(P_0)}{P_0P} = \frac{h}{P_0P}$$

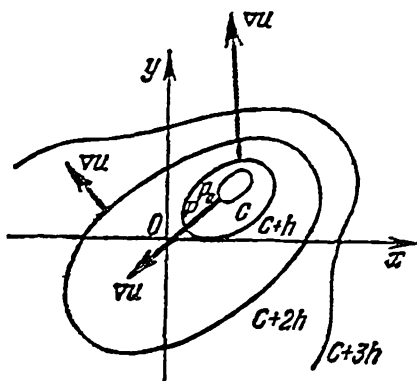


Fig. 157

where $P_0(x_0, y_0)$ is the given point of the level line $u(x, y) = C$ ($C = u(x_0, y_0)$) and P is the

corresponding point of the level line $u(x, y) = C + h$. The magnitude of h is given, and the length of the line segment P_0P can be measured as the distance along the normal between the neighbouring level lines depicted in the diagram. Thus we find an approximation to the modulus of the gradient since it is equal to the derivative in the direction of the gradient:

$$|\text{grad } u| \approx \frac{h}{P_0P}$$

The construction of the vector $\text{grad } u$ according to the scheme described here is shown in Fig. 157.

QUESTIONS

1. What do we call a function of two independent variables? What is the domain of definition of such a function?
2. Describe the methods of tabular and analytical specification on a function of two independent variables.
3. What do we call the graph of a function of two independent variables? Describe the method of graphical representation of such a function.
4. Describe the method of investigation of a function of two independent variables with the aid of functions of one independent variable. What is a level line of a function $z = f(x, y)$?
5. What do we call the limit of a function $z = f(x, y)$ as $x \rightarrow x_0$ and $y \rightarrow y_0$?
6. State the definition of continuity for a function of two independent variables at a point and in a domain. Give examples of discontinuous functions.
7. State the definition of the partial derivative of a function of two independent variables with respect to one of them. Generalize this definition to functions of more than two independent variables.
8. What is the geometrical meaning of the partial derivatives of a function $z = f(x, y)$ in Cartesian coordinates?
9. What do we call the partial increment and the partial differential of a function $z = f(x, y)$ with respect to x ? How is a partial differential of a function expressed in terms of the corresponding partial derivative?
10. What is the geometrical meaning of the partial differentials of a function $z = f(x, y)$ in Cartesian coordinates?
11. What do we call the total increment and the total differential of a function $z = f(x, y)$? How is the total differential of a function expressed in terms of its partial derivatives?
- 12*. What is a differentiable function of several independent variables? State and prove the sufficient condition for differentiability.
13. What do we call the tangent plane to a surface at its given point? Derive the equation of the tangent plane.
14. What is the geometrical meaning of the total differential of a function $z = f(x, y)$ in Cartesian coordinates?
15. How is the total differential of a function applied to approximate calculation of its values?
16. Derive the formula for the limiting absolute error in the computation of the value of a function of several variables.
17. Derive the formulas for the limiting relative errors occurring in the computation of a product and a quotient.
18. State the definition of a partial derivative of the n th order of a function of two independent variables.

19. State the theorem on the equality of the mixed partial derivatives of the second order.

20. Extend the definition of a partial derivative of the n th order to functions of several independent variables.

21. State the definition of the total differential of the second order of a function of two arguments and write down the formula expressing the second differential.

22. State the necessary and sufficient conditions for an expression $P(x, y)dx + Q(x, y)dy$ to be a total differential. Prove these conditions. State the corresponding conditions for functions of three variables.

23. Derive the differentiation rule for a composite function.

24. What do we call the total derivative?

25. Describe the invariance property of the form of the total differential of a function of several variables.

26. Enumerate the basic rules for finding the differential of a function of any number of independent variables.

27. State the existence theorem for an implicit function.

28. State the differentiation rule for an implicit function.

29. Write down the equations of the tangent plane and of the normal to a surface specified by an equation of the general form in Cartesian coordinates.

30. How can we form the equations of the tangent line and of the normal plane to a curve specified as a line of intersection of two surfaces?

31. State the definition of an extremum (a maximum or a minimum) of a function of two variables.

32. What are the necessary conditions for extremum of a function of two independent variables? Describe the geometrical meaning of these conditions.

33. State the sufficient conditions for extremum for a function of two independent variables.

34. Describe the method for determining the greatest and the least values of a function of two independent variables defined in a closed domain.

35*. State the definition of the point of a conditional extremum of a function $f(x, y)$ on a line L .

36*. Describe the techniques for finding the points of conditional extremum.

37*. State the general conditional extremum problem for functions of three and more independent variables.

38. What do we call a scalar field? What is a level surface of a scalar field?

39. Derive the formula for the directional derivative.

40. State the definition of the gradient of a scalar field.

41. State and prove the basic properties of the gradient.

Chapter VIII

DOUBLE AND TRIPLE INTEGRALS

§ 1. Double Integrals

128. The Volume of a Curvilinear Cylinder. The Double Integral. Now we proceed to study integral calculus for functions of two independent variables. To begin with, we shall consider the concrete problem of computing the area of *curvilinear cylinder*, which leads directly to the notion of the double integral.

A *curvilinear cylinder (cylindroid)* is a solid with boundary consisting of a closed domain D of the xy -plane, a surface $z=f(x, y)$ where $f(x, y)$ is a continuous and nonnegative function in the domain D and the cylindrical surface whose elements (generators) are parallel to Oz and directrix is the boundary of the domain D (see Fig. 158).

The domain D is called the *base* of the cylindroid; it is supposed to be bounded, and its boundary usually consists of one or several closed piece-wise smooth curves. In a particular case there may be no lateral cylindrical surface; an instance of such a cylindroid is the solid bounded by the plane Oxy and the upper hemisphere $z=\sqrt{R^2-x^2-y^2}$.

The volume of an arbitrary solid can usually be represented as a sum or difference of the volumes of cylindrical bodies.

When defining the *volume of a solid* we proceed from the following two fundamental hypotheses:

(1) If a body is broken up into parts its volume is equal to the sum of the volumes of the parts.

(2) The volume of a right cylinder, that is, of a cylindroid (with element parallel to Oz) bounded by the xy -plane and another plane parallel to the former, is equal to the product of the base area by the altitude of the body.

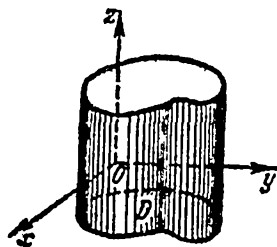


Fig. 158

Let us denote the sought-for volume by V . Now we break up the base of the curvilinear cylinder (i.e. the domain D) into n closed subdomains (cells) of arbitrary shape. This is referred to as a *partition* of the domain D . Let the cells be numbered in an arbitrary order and denoted $\sigma_1, \sigma_2, \dots, \sigma_n$, their areas being,

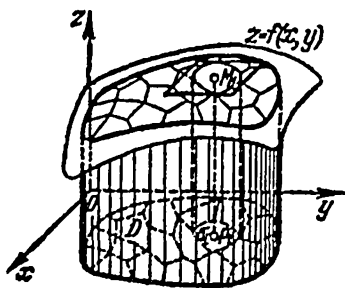


Fig. 159

respectively, $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$. Next we draw a cylindrical surface through the boundary of each subdomain (these boundaries serve as the directrices of the cylinders) with elements parallel to Oz . These cylindrical surfaces divide the upper bounding surface (with equation $z=f(x, y)$) into n parts, and the whole cylindroid is thus broken up into n subcylindroids (Fig. 159). Let us choose an arbitrary point $P_i(x_i, y_i)$ in each cell σ_i and replace the corresponding subcylindroid by the right cy-

linder with the same base and altitude equal to $z_i=f(x_i, y_i)$. This results in an n -step solid whose volume is

$$\begin{aligned} V_n &= f(x_1, y_1) \Delta\sigma_1 + f(x_2, y_2) \Delta\sigma_2 + \dots + f(x_n, y_n) \Delta\sigma_n = \\ &= \sum_{i=1}^n f(x_i, y_i) \Delta\sigma_i \end{aligned}$$

It appears natural to regard the volume V_n of the n -step body thus constructed as an approximation to the volume V of the given curvilinear cylinder and assume that, the greater the number n and the smaller each of the cells, the more accurate is the approximation. The next step is to pass to the limit as $n \rightarrow \infty$ on condition that not only the area of each subdomain tends to zero but also all its dimensions. To state what has been said in a more precise manner we introduce the notion of the *diameter of a bounded closed domain as the greatest distance between the points of its boundary*; the above condition means that the diameters of the cells should tend to zero as $n \rightarrow \infty$; then the cells themselves contract to points.¹

Therefore, in accordance with what has been said, we define the sought-for volume V as the limit of V_n when the *maximal diame-*

¹ For example, the diameter of a rectangle is equal to the length of its diagonal and that of an ellipse to its major axis; for a circle this definition of the diameter coincides with the ordinary one. If it is only required that the area of a domain tends to zero the domain itself does not necessarily contract to a point. For instance, the area of the rectangle with *constant* base and infinitesimal altitude tends to zero but the rectangle itself contracts not to a point but to its base, that is, to a line segment.

ter of the partition (that is, the greatest of the diameters of the subdomains the original domain is broken into) tends to zero while $n \rightarrow \infty$:

$$V = \lim V_n = \lim \sum_{i=1}^n f(x_i, y_i) \Delta \sigma_i$$

(The maximal diameter of a partition of a domain is sometimes called the *fineness* of the partition.)

There are various problems leading to the computation of the limit of a sum of this kind for functions of two variables.

Now we consider this question in the general form. Let $f(x, y)$ be a function of two variables (not necessarily positive) bounded in a closed finite domain D . Let us divide the domain D into cells as was indicated above, choose an arbitrary point $P_i(x_i, y_i)$ in each cell and compose the sum

$$\sum_{i=1}^n f(x_i, y_i) \Delta \sigma_i \quad (*)$$

where $f(x_i, y_i)$ is the value of the function $f(x, y)$ at the point P_i and $\Delta \sigma_i$ the area of the subdomain σ_i . A sum of form $(*)$ is called an *integral sum associated with the function $f(x, y)$ and the domain D for the given partition of the domain into n subdomains*.

Definition. The limit¹ of integral sum $(*)$ (provided it exists) as the fineness of the partition tends to zero is called the *double integral of the function $f(x, y)$ over the domain D* .

This is written as

$$\lim \sum_{i=1}^n f(x_i, y_i) \Delta \sigma_i = \iint_D f(x, y) d\sigma$$

(read "the double integral of $f(x, y) d\sigma$ over the domain D "). The expression $f(x, y) d\sigma$ symbolizing the form of the summands entering into the integral sum is termed the *element of integration*, the function $f(x, y)$ is called the *integrand*, $d\sigma$ is the *element of area* and D the *domain of integration*, the arguments x and y being referred to as the *variables of integration*.

Thus, we can say that the volume of a curvilinear cylinder considered above is expressed by the double integral of the function $f(x, y)$ over the domain serving as the base of the cylinder:

$$V = \iint_D f(x, y) d\sigma$$

¹ The precise definition of the limit of integral sum $(*)$ is stated by analogy with the corresponding definition given in Sec. 87.

There holds the following theorem analogous to the existence theorem for the ordinary definite integral (Sec. 87):

Theorem (On the Existence of the Double Integral). If $f(x, y)$ is a continuous function in a bounded closed domain D its integral sum tends to a definite limit as the maximal diameter of the partition tends to zero. This limit, that is, the double integral $\iint_D f(x, y) d\sigma$, is independent of the way in which the domain D is divided into the cells σ_i and on the choice of the points P_i in the cells.

The double integral is of course a number dependent solely on the integrand and on the domain of integration and completely independent of the notation of the variables of integration; for instance,

$$\iint_D f(x, y) d\sigma = \iint_D f(u, v) d\sigma$$

As will be shown later, the computation of a double integral can usually be reduced to two ordinary definite integrations.

129. Properties of Double Integrals. The reader has undoubtedly noted that the fundamental idea of the construction of the double integral is analogous to that of the definite integral of a function of one independent variable. Therefore the properties of the double integral (and also the proofs of these properties) are almost the same as in the case of the ordinary definite integral. That is why, assuming that all the functions involved are continuous and their double integrals thus exist, we shall only state these properties without presenting the proofs.

I. The double integral of a sum of a finite number of functions is equal to the sum of the double integrals of the summands:

$$\begin{aligned} & \iint_D [f(x, y) + \varphi(x, y) + \dots + \psi(x, y)] d\sigma = \\ &= \iint_D f(x, y) d\sigma + \iint_D \varphi(x, y) d\sigma + \dots + \iint_D \psi(x, y) d\sigma \end{aligned}$$

II. A constant factor in the integrand can be taken outside the symbol of the double integral:

$$\iint_D cf(x, y) d\sigma = c \iint_D f(x, y) d\sigma$$

III. If the domain of integration D is split into two domains D_1 and D_2 having no interior points in common then

$$\iint_D f(x, y) d\sigma = \iint_{D_1} f(x, y) d\sigma + \iint_{D_2} f(x, y) d\sigma$$

IV. If two functions $f(x, y)$ and $\varphi(x, y)$ satisfy the condition $f(x, y) \geq \varphi(x, y)$ at all the points of the domain of integration D then

$$\iint_D f(x, y) d\sigma \geq \iint_D \varphi(x, y) d\sigma$$

These integrals are only equal when both functions coincide.

In particular, Property IV implies that

if the integrand does not change sign in the domain of integration the double integral is a number of the same sign as the integrand function.

Property III and the above consequence of Property IV lead to a more precise geometrical interpretation of the double integral. Let us agree to assign the plus sign to the volume of a cylindroid lying *above* the plane Oxy and the minus sign to that *below* Oxy . Then, if $z = f(x, y)$ is the equation of the bounding surface the double integral $\iint_D f(x, y) d\sigma$ is equal to the *algebraic sum* of the volumes of the bodies corresponding to the positive and negative values of the function $f(x, y)$.

If the integrand in a double integral is identically equal to 1 the integral is equal to the area S of the domain of integration, i.e.

$$\iint_D d\sigma = S$$

In this case the double integral expresses the volume of the right cylinder with base D and unit altitude; it is clear that the numerical value of this volume is equal to that of the area of the base of the cylinder.

V. If m and M are the least and the greatest values of the integrand in the domain of integration D the value of the double integral lies between the products of the least and the greatest values m and M by the area of the domain of integration:

$$mS \leq \iint_D f(x, y) d\sigma \leq MS$$

where S is the area of the domain D .

VI. The double integral is equal to the product of the value of the integrand at a point of the domain of integration by the area of the domain of integration, that is,

$$\iint_D f(x, y) d\sigma = f(\xi, \eta) S$$

The value $f(\xi, \eta)$ entering into the latter relation is called the *mean value of the function $f(x, y)$ in the domain D* (cf. Sec. 90).

The geometrical meaning of the last property (known as the *mean value theorem for the double integral*) is that there exists a right cylinder whose base coincides with the base D of the given curvilinear cylinder and whose altitude is equal to the z -coordinate of the surface $z=f(x, y)$ at a point (ξ, η) of the base such that its volume is equal to that of the given cylindroid. This z -coordinate is nothing but the *mean value* of the function $f(x, y)$ in the domain D .

130. Evaluating Double Integrals. To evaluate the double integral $\iint_D f(x, y) d\sigma$ it is convenient to use a special representation of the element of area $d\sigma$. Let us divide the domain of integration D lying in the xy -plane into subdomains by means of two systems of *coordinate lines*: $x=\text{const}$ and $y=\text{const}$. These are straight lines parallel, respectively, to Oy and Ox , the subdomains (cells) being rectangles with sides parallel to the coordinate axes. It is clear that the area of each cell not adjoining the boundary of the domain D is equal to the product of its dimensions Δx and Δy . Therefore the element of area $d\sigma$ can be written as

$$d\sigma = dx dy$$

that is, *the element of area in Cartesian coordinates is equal to the product of the differentials of the independent variables.*¹ Hence, we have

$$\iint_D f(x, y) d\sigma = \iint_D f(x, y) dx dy \quad (*)$$

In the computation of double integral $(*)$ we shall take advantage of the fact that the integral expresses the volume V of the curvilinear cylinder with base D bounded by the surface $z=f(x, y)$. We remind the reader that the problem of computing the volume of a solid was already considered in connection with the applications of the definite integral to geometrical problems. In Sec. 101 we derived the formula

$$V = \int_a^b S(x) dx \quad (**)$$

for the volume of a solid where $S(x)$ is the cross-section area in the plane perpendicular to the axis of abscissas cutting that axis at the point x while $x=a$ and $x=b$ are the equations of the planes bounding the given solid (see Fig. 122 on p. 348). We

¹ A cell adjoining the boundary of the domain of integration is not necessarily of the shape of a rectangle, and its area may differ from $dx dy$. It can be proved that the part of the integral sum including such subdomains vanishes in the limiting process.

shall apply this formula to evaluating the double integral

$$\int_D \int f(x, y) dx dy$$

Let us begin with the case when the domain of integration D satisfies the following condition: every straight line parallel to the x -axis or to the y -axis and cutting the boundary of the domain has at most two common points with it. A cylindrical body of this type is depicted in Fig. 160.

Let us include the domain D in the rectangle

$$a \leq x \leq b, \quad c \leq y \leq d$$

whose sides touch the boundary of the domain at points A, B, C and E . The interval $[a, b]$ is the orthogonal projection of the

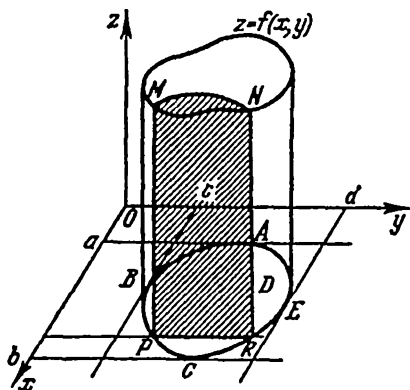


Fig. 160

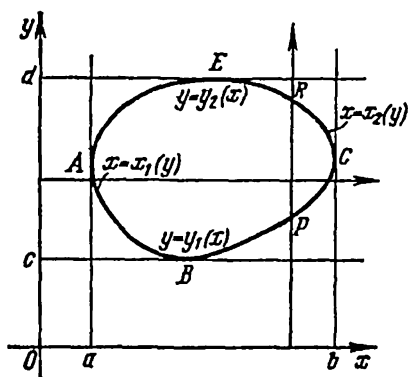


Fig. 161

domain D on Ox and the interval $[c, d]$ the orthogonal projection of D on Oy . Fig. 161 shows the xy -plane and the domain D lying in it.

The points A and C divide the boundary of the domain D into two parts: the arc ABC and the arc AEC ; each of them has exactly one common point with any straight line parallel to the y -axis and cutting the domain D . Therefore the equations of these arcs can be written in the form solved with respect to y ; let

$y = y_1(x)$ be the equation of ABC

and

$y = y_2(x)$ be the equation of AEC

Similarly, the points B and E break up the boundary into two arcs BAE and BCE whose equations can be written as

$$x = x_1(y) \text{ (the equation of } BAE)$$

and

$$x = x_2(y) \text{ (the equation of } BCE)$$

Let us consider the section of the cylindroid by an arbitrary plane parallel to the plane Oyz ; the equation of such a plane is $x = \text{const}$, $a \leq x \leq b$ (see Fig. 160). In the section we obtain a curvilinear trapezoid $PMNR$ whose area is expressed by the integral of $f(x, y)$ regarded as a function of one variable y ranging from the ordinate of the point P to the ordinate of the point R . P is the lowest (in the xy -plane) common point of the straight line $x = \text{const}$ with the domain D and R is its highest common point with D . From the equations of the curves ABC and AEC it follows that, for the given value of x , the ordinates of these points are, respectively, equal to $y_1(x)$ and $y_2(x)$.

Consequently, the integral

$$\int_{y_1(x)}^{y_2(x)} f(x, y) dy$$

expresses the area of the plane section $PMNR$. It is clear that the value of this integral depends on the chosen value of x ; in other words, the area of this section is a function of x ; we shall denote it $S(x)$:

$$S(x) = \int_{y_1(x)}^{y_2(x)} f(x, y) dy$$

(It can be proved that if the functions $f(x, y)$, $y_1(x)$ and $y_2(x)$ are continuous the function $S(x)$ is continuous as well.)

According to formula (**), the volume of the entire curvilinear cylinder is equal to the integral of $S(x)$ over the interval of variation of x , that is, over $[a, b]$. Substituting the expression of $S(x)$ into that integral we finally receive

$$V = \int_D \int f(x, y) dx dy = \int_a^b \left[\int_{y_1(x)}^{y_2(x)} f(x, y) dy \right] dx$$

or, in a more convenient form¹,

$$\int_D \int f(x, y) dx dy = \int_a^b dx \int_{y_1(x)}^{y_2(x)} f(x, y) dy \quad (A)$$

The limits of integration in the inner integral in formula (A) are variable; they indicate the range of variation of y for a

¹ The derivation of formula (A) presented here is based on visual geometrical considerations and is not of course a rigorous analytical proof.

constant value of the other argument x . The limits of integration in the outer integral are constant; they indicate the range of the variable x .

Interchanging the roles of x and y , that is, starting from the sections of the body by the planes $y = \text{const}$ ($c \leq y \leq d$), we first find that the area $Q(y)$ of such a section is equal to $\int_{x_1(y)}^{x_2(y)} f(x, y) dx$ where y is considered constant in the integration process and then integrate the function $Q(y)$ between the limits of variation of y , i.e. from c to d , to obtain the alternative expression of the double integral:

$$\iint_D f(x, y) dx dy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} f(x, y) dx \quad (B)$$

Here the inner integration is performed with respect to x and the outer with respect to y .

Formulas (A) and (B) show that the computation of the double integral reduces to two consecutive ordinary definite integrations; one should memorize that in the inner integral one of the variables is considered constant in the integration process. The expressions on the right-hand sides of formulas (A) and (B) are called (*twofold*) *iterated* or *repeated integrals*, the whole computation process being referred to as the *reduction of the double integral to an iterated integral*.

The reduction of a double integral to an iterated integral becomes particularly simple when the domain of integration D is a rectangle with sides parallel to the coordinate axes (see Fig. 162). In this case not only the limits of integration in the outer integral are constant but those in the inner integral as well:

$$\iint_D f(x, y) d\sigma = \int_a^b dx \int_c^d f(x, y) dy = \int_c^d dy \int_a^b f(x, y) dx$$

In more complicated cases to reduce a double integral to an iterated one we should construct the domain of integration (it is advisable to represent it in the xy -plane as was done in Fig. 161), decide on the order of integrations, that is, choose one of the variables for the inner integration and the other for the outer integration, and then set up the limits of integration.

In the examples below we demonstrate the technique of setting the limits of integration.

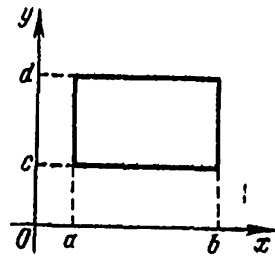


Fig. 162

Examples. (1) Let us reduce the double integral $\iint_D f(x, y) dx dy$ taken over the triangular domain D bounded by the straight lines $y=0$, $y=x$ and $x=a$ (Fig. 163) to an iterated integral. If we first integrate with respect to y and then with respect to x the inner integration is performed from the line $y=0$ to the line $y=x$ while the outer integration from the point $x=0$ to the point $x=a$. Thus,

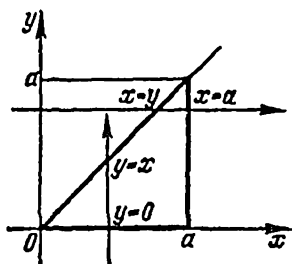


Fig. 163

$$\iint_D f(x, y) dx dy = \int_0^a dx \int_0^x f(x, y) dy$$

Reversing the order of integration we receive

$$\iint_D f(x, y) dx dy = \int_0^a dy \int_y^a f(x, y) dx$$

(2) Let us reduce the integral $\iint_D f(x, y) dx dy$ to an iterated integral for the domain D bounded by the lines $y=0$, $y=x^2$ and $x+y=2$.

The domain D is shown in Fig. 164 where the coordinates of its extreme points are written down. The shape of the domain indicates that it is more convenient to integrate first with respect to x and then with respect to y :

$$\iint_D f(x, y) dx dy = \int_0^1 dy \int_{\sqrt{y}}^{2-y} f(x, y) dx$$

If the order of integration is reversed the result cannot be written in the form of one iterated integral since different parts of the line OBA are described by different equations. Therefore we split the domain D into two: OBC and CBA , which results in

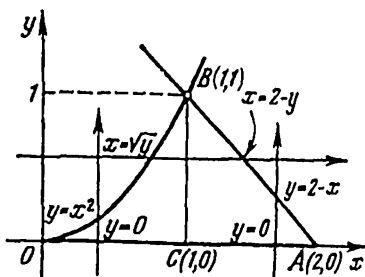


Fig. 164

$$\iint_D f(x, y) dx dy = \int_0^1 dx \int_0^{x^2} f(x, y) dy + \int_1^2 dx \int_0^{2-x} f(x, y) dy$$

This example shows that it is extremely important to make an appropriate choice of the order of integration (in this example we obtain a more complicated expression in the latter case than in the former).

Formulas (A) and (B) can sometimes be applied to domains of a more general form. For instance, formula (A) can be used in

the case shown in Fig. 165 and formula (B) in the case in Fig. 166. As a rule, complicated domains can be split into a finite number of simpler domains (e.g. see Fig. 167) admitting the application of formulas (A) and (B) to the computation of the double integrals over these subdomains, which results in the computation of the double integral over the original domain.

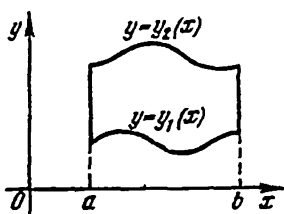


Fig. 165

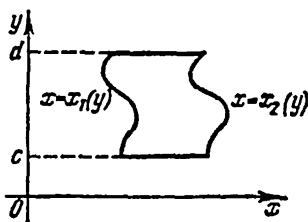


Fig. 166

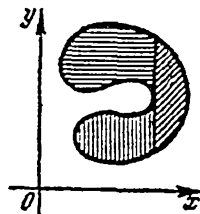


Fig. 167

Let us consider some examples demonstrating the evaluation of double integrals.

Examples. (1) Let us evaluate the double integral of the function

$$z = 1 - \frac{1}{3}x - \frac{1}{4}y$$

over the rectangle D determined by the conditions $-1 \leq x \leq 1$, $-2 \leq y \leq 2$:

$$I = \iint_D \left(1 - \frac{1}{3}x - \frac{1}{4}y\right) dx dy$$

Geometrically, the integral I is equal to the volume of the right quadrangular prism with the domain D as base truncated by the plane

$$\frac{1}{3}x + \frac{1}{4}y + z = 1$$

(see Fig. 168). Let us take the iterated integral with inner integration with respect to y and outer integration with respect to x :

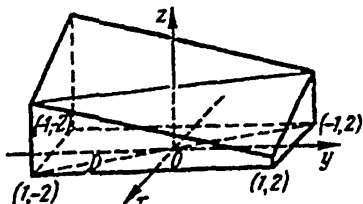


Fig. 168

$$\begin{aligned} I &= \int_{-1}^1 dx \int_{-2}^2 \left(1 - \frac{1}{3}x - \frac{1}{4}y\right) dy = \int_{-1}^1 \left(y - \frac{1}{3}xy - \frac{1}{8}y^2\right) \Big|_{-2}^2 dx = \\ &= \int_{-1}^1 \left(4 - \frac{4}{3}x\right) dx = \left(4x - \frac{2}{3}x^2\right) \Big|_{-1}^1 = 8 \end{aligned}$$

The same result is obtained when the order of integration is reversed:

$$\begin{aligned} I &= \int_{-2}^2 dy \int_{-1}^1 \left(1 - \frac{1}{3}x - \frac{1}{4}y\right) dx = \int_{-2}^2 \left(x - \frac{1}{6}x^2 - \frac{1}{4}yx\right) \Big|_{-1}^1 dy = \\ &= \int_{-2}^2 \left(2 - \frac{1}{2}y\right) dy = \left(2y - \frac{1}{4}y^2\right) \Big|_{-2}^2 = 8 \end{aligned}$$

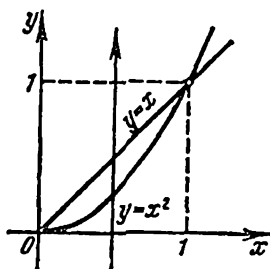


Fig. 169

(2) Let us compute the double integral

$$\iint_D (x+y) dx dy$$

over the domain D bounded by the lines $y=x$ and $y=x^2$. The domain D is shown in Fig. 169. Performing the inner integration with respect to y and the outer integration with respect to x we obtain

$$\begin{aligned} \iint_D (x+y) dx dy &= \int_0^1 dx \int_{x^2}^x (x+y) dy = \int_0^1 \left[xy + \frac{y^2}{2}\right]_{x^2}^x dx = \\ &= \int_0^1 \left(x^2 + \frac{x^2}{2} - x^3 - \frac{x^4}{2}\right) dx = \left(\frac{x^3}{2} - \frac{x^4}{4} - \frac{x^5}{10}\right) \Big|_0^1 = \frac{3}{20} \end{aligned}$$

The correctness of the result is readily checked if we change the order of integration:

$$\begin{aligned} \iint_D (x+y) dx dy &= \int_0^1 dy \int_y^{\sqrt{y}} (x+y) dx = \int_0^1 \left[\frac{x^2}{2} + xy\right]_y^{\sqrt{y}} dy = \\ &= \int_0^1 \left(\frac{y}{2} + y\sqrt{y} - \frac{y^2}{2} - y^2\right) dy = \left(\frac{y^2}{4} + \frac{2y^{\frac{5}{2}}}{5} - \frac{y^3}{2}\right) \Big|_0^1 = \frac{3}{20} \end{aligned}$$

(3) Let us compute the volume of the solid bounded by the cylindrical surfaces $z=4-y^2$ and $y=\frac{x^2}{2}$ and the plane $z=0$ (Fig. 170a).

The upper boundary of the solid is the surface with equation $z=4-y^2$. The domain of integration D is bounded by the parabola $y=\frac{x^2}{2}$ and the line of intersection of the cylinder $z=4-y^2$ with the plane $z=0$, i.e. the straight line $y=2$ (see Fig. 170b).

The solid being symmetric about the yz -plane, it is sufficient to compute half the sought-for volume:

$$\begin{aligned}\frac{1}{2}V &= \int_0^2 dx \int_{\frac{x^2}{2}}^2 (4-y^2) dy = \int_0^2 \left[4y - \frac{y^3}{3} \right]_{\frac{x^2}{2}}^2 dx = \\ &= \int_0^2 \left(8 - \frac{8}{3} - 2x^2 + \frac{x^6}{24} \right) dx = \left(\frac{16}{3}x - \frac{2x^3}{3} + \frac{x^7}{168} \right)_0^2 = \frac{128}{21}\end{aligned}$$

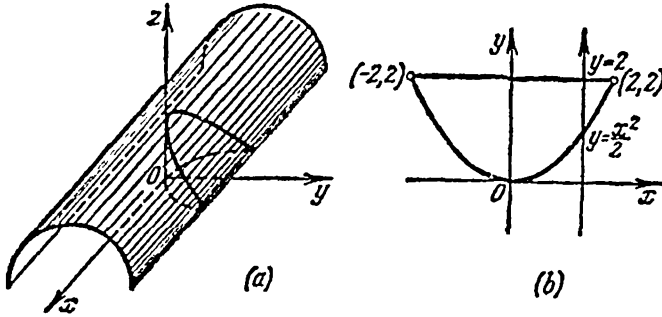


Fig. 170

Consequently,

$$V = \frac{256}{21} \approx 12.2$$

Let the reader compute this integral for the reversed order of integration.

(4) Let us find the volume V of the solid bounded by the surface

$$z = 1 - 4x^2 - y^2$$

and by the plane Oxy .

The solid in question is a segment of the elliptical paraboloid lying above the plane Oxy (Fig. 171). The paraboloid cuts the xy -plane along the ellipse

$$4x^2 + y^2 = 1$$

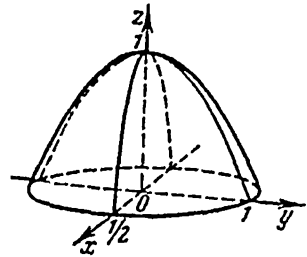


Fig. 171

The problem thus reduces to computing the volume of the "cylindroid" without lateral cylindrical surface bounded above by the paraboloid

$$z = 1 - 4x^2 - y^2$$

and having the interior of the ellipse as base.

The solid under consideration being symmetric with respect to the planes Oxz and Oyz , it is sufficient to determine the quarter volume lying in the first coordinate trihedral. The latter is equal to the double integral over the domain specified by the conditions $4x^2 + y^2 \leq 1$, $x \geq 0$, and $y \geq 0$, i.e. over the quarter ellipse. Integrating with respect to y and then with respect to x we receive

$$\frac{1}{4} V = \int_0^{\frac{1}{2}} dx \int_0^{\sqrt{1-4x^2}} (1-4x^2-y^2) dy = \frac{2}{3} \int_0^{\frac{1}{2}} (1-4x^2)^{\frac{3}{2}} dx$$

The substitution $2x = \sin t$ results in

$$\frac{1}{4} V = \frac{2}{3} \cdot \frac{1}{2} \int_0^{\frac{\pi}{2}} \cos^4 t dt = \frac{2}{3} \cdot \frac{1}{2} \cdot \frac{3}{16} \pi$$

(see Sec. 94) whence

$$V = \frac{\pi}{4}$$

131. Change of Variables in the Double Integral. Polar Coordinates.

I. The rule for change of variable in the definite integral (see Sec. 94, II) plays an extremely important role in practical calculations; it states that

$$\int_{x_1}^{x_2} f(x) dx = \int_{u_1}^{u_2} f[\psi(u)] \psi'(u) du$$

The function $x = \psi(u)$ is usually monotone; in the case of monotonicity it specifies a one-to-one correspondence between the points of the interval $[u_1, u_2]$ of variation of the variable u and the points of the interval $[x_1, x_2]$ of variation of the variable x (see Sec. 18).

Thus, when performing the change of variable according to the formula $x = \psi(u)$ we must replace dx by $dx = \psi'(u) du$ and the old limits of integration x_1 and x_2 for the variable x by the new ones u_1 and u_2 for the variable u .

The rule for change of variables in the double integral is essentially more complicated and its detailed discussion falls out of the framework of our course. We shall confine ourselves to the final formula and to the example of its application in the case of the transformation from Cartesian coordinates to polar ones.

If the variables x and y in the double integral are replaced by some new variables u and v given by formulas

$$x = x(u, v), \quad y = y(u, v) \quad (*)$$

the formula for change of variables has the form

$$\iint_D f(x, y) d\sigma = \iint_{D^*} f[x(u, v), y(u, v)] \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv \quad (**)$$

where

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u}$$

The latter expression is (Jacobi's)¹ functional determinant also termed a *Jacobian* whose elements are the partial derivatives of the functions (*).

Thus, when making a change of variables specified by formula (*) we should replace $d\sigma = dx dy$ according to the formula

$$d\sigma = \left| \frac{\partial(x, y)}{\partial(u, v)} \right| du dv$$

and also replace the original domain of integration D with respect to the variables x and y by the corresponding new domain of integration D^* with respect to the variables u and v . The new expression for $d\sigma$ is the *element of area in the coordinates u and v* .²

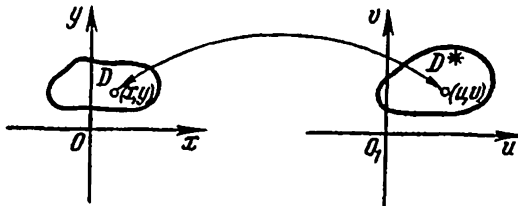


Fig. 172

We shall not present the detailed statement of all conditions providing the validity of formula (**) and only note that the functions (*) are supposed to be continuous together with their partial derivatives and that the correspondence between the domains D and D^* should be one-to-one at least for the interior points (see Fig. 172).

An appropriate change of variables may essentially simplify the given integral, for instance, it may lead to constant limits of integration in the transformed integral.

¹ Jacobi, K.G.J. (1804-1851), a German mathematician.

² The new variables u and v themselves are called *curvilinear coordinates*. There are various systems of curvilinear coordinates playing an important role in mathematics and its applications.

II. The double integral in polar coordinates. Let us apply the general formula (**) to the transformation from Cartesian coordinates (x and y) to polar coordinates (which we denote r and φ instead of u and v):

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

It is assumed that $r \geq 0$ and that the angle φ ranges from 0 to 2π or from $-\pi$ to π .

The Jacobian of this transformation (or, synonymously, of this mapping) is

$$\frac{\partial(x, y)}{\partial(r, \varphi)} = \frac{\partial x}{\partial r} \frac{\partial y}{\partial \varphi} - \frac{\partial x}{\partial \varphi} \frac{\partial y}{\partial r} = \cos \varphi \cdot r \cos \varphi - (-r \sin \varphi) \sin \varphi = r$$

Therefore,

$$\iint_D f(x, y) dx dy = \iint_{D^*} f(r \cos \varphi, r \sin \varphi) r d\varphi dr \quad (***)$$

where D and D^* are the domains in the planes Oxy and $O_1r\varphi$ corresponding to each other under this transformation (when speaking about the plane $O_1r\varphi$ we interpret r and φ as Cartesian coordinates in an auxiliary plane).

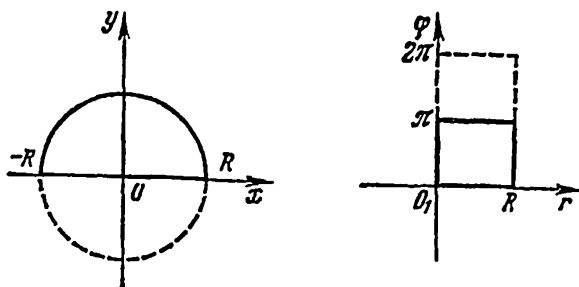


Fig. 173

For example, let D be a semicircle of radius R with centre at the origin lying in the half-plane $y \geq 0$ (Fig. 173). The corresponding domain D^* in the auxiliary plane $O_1r\varphi$ is the rectangle $0 \leq r \leq R$, $0 \leq \varphi \leq \pi$.¹ Hence

$$\iint_D f(x, y) dx dy = \int_0^\pi d\varphi \int_0^R f(r \cos \varphi, r \sin \varphi) r dr$$

If D is the whole circle of radius R with centre at the origin the corresponding domain D^* is formed of the two rectangles shown

¹ In this transformation the point $(0, 0)$ of the xy -plane goes into the line segment $[0, \pi]$ on the axis $O_1\varphi$ in the $r\varphi$ -plane. This violation of the one-to-one correspondence occurs on the boundary of the domain D^* , and the transformation formula for the integral remains valid.

in Fig. 173 in heavy and dotted lines; in this case the integration with respect to φ goes from 0 to 2π .

Let the reader plot the domains D^* corresponding to the domains D of the shape of a sector of a circle with centre at the origin and of an annulus between concentric circles.

The formula for the element of area in polar coordinates can be obtained directly on the basis of visual geometrical considerations. To this end, let us construct the coordinate lines $r = \text{const}$ and $\varphi = \text{const}$ (of the polar coordinate system r, φ) lying in the xy -plane; they break up the plane Oxy into curvilinear quadrilaterals bounded by arcs of concentric circles and their radii (see Fig. 174). Consider the quadrilateral given in heavy line in Fig. 174. Its area $\Delta\sigma$ is expressed as

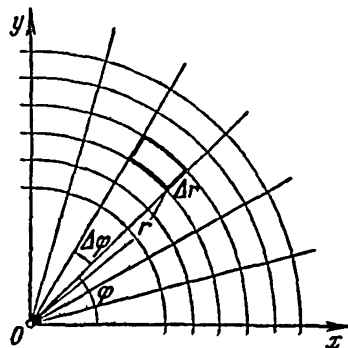


Fig. 174

$$\Delta\sigma = \frac{1}{2}(r + \Delta r)^2 \Delta\varphi - \frac{1}{2}r^2 \Delta\varphi = r \Delta r \Delta\varphi + \frac{1}{2}\Delta r^2 \Delta\varphi$$

The second summand $\frac{1}{2}\Delta r^2 \Delta\varphi$ on the right is an infinitesimal of higher order than the first summand; on discarding this term we arrive at the approximate relation

$$\Delta\sigma \approx r \Delta r \Delta\varphi$$

which shows that it is natural to expect that the element of area in polar coordinates is written as $d\sigma = r dr d\varphi$. Thus, we again come to formula (**).

To reduce a double integral written in polar coordinates to an iterated integral it is often possible to resort to the rules below without constructing the domain D^* in the auxiliary plane $O_1r\varphi$.

1. Let the pole lie outside the domain of integration D enclosed within two rays $\varphi = \varphi_1$ and $\varphi = \varphi_2$, and possessing the property that any straight line $\varphi = \text{const}$ has at most two common points with its boundary (see Fig. 175a; a domain of this type can also be of the shape shown in Fig. 175b).

Let the polar equations of the arcs ADC and ABC be, respectively, $r = r_1(\varphi)$ and $r = r_2(\varphi)$, both functions being continuous in the closed interval $[\varphi_1, \varphi_2]$.

Performing the inner integration with respect to r between the limits of its variation for a constant φ , that is, from $r_1(\varphi)$ to $r_2(\varphi)$, and the outer integration with respect to φ from φ_1 to φ_2 ,

we receive

$$I = \int_{\varphi_1}^{\varphi_2} d\varphi \int_{r_1(\varphi)}^{r_2(\varphi)} f(r \cos \varphi, r \sin \varphi) r dr$$

The opposite order of integration, that is, the inner integration with respect to φ and the outer with respect to r , is encountered more rarely. If the line ADC in Fig. 175a degenerates into the point O we have $r_1(\varphi) = 0$.

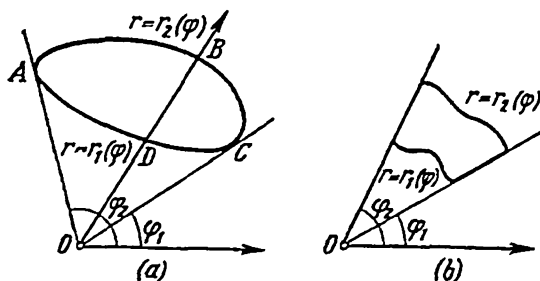


Fig. 175

In the particular case when the domain of integration is a part of an annulus between concentric circles determined by conditions $r_1 \leq r \leq r_2$, $\varphi_1 \leq \varphi \leq \varphi_2$ the limits of integration with respect to both variables are constant:

$$I = \int_{\varphi_1}^{\varphi_2} d\varphi \int_{r_1}^{r_2} f(r \cos \varphi, r \sin \varphi) r dr$$

2. Let the pole be inside the domain of integration, the latter being such that any polar radius meets its boundary at one point (such a domain is said to be *starlike* (or *starshaped*) with respect to the pole). Performing the inner integration with respect to r and the outer with respect to φ we obtain

$$I = \int_0^{2\pi} d\varphi \int_0^{r(\varphi)} f(r \cos \varphi, r \sin \varphi) r dr$$

where $r = r(\varphi)$ is the polar equation of the boundary of the domain.

In particular, for $r = r(\varphi) = R = \text{const}$, that is, when the domain of integration is the circle of radius R and centre at the pole, we have

$$I = \int_0^{2\pi} d\varphi \int_0^R f(r \cos \varphi, r \sin \varphi) r dr$$

Examples. (1) Let us set up the limits of integration in the double integral over the domain D of the shape of the disc $x^2 + y^2 \leq ax$ (Fig. 176). Passing to polar coordinates we write down the equation of the circle in the form $r = a \cos \varphi$. Here we have $r_1(\varphi) = 0$ and $r_2(\varphi) = a \cos \varphi$. The limits of variation of φ are $-\frac{\pi}{2}$ and $\frac{\pi}{2}$. Therefore,

$$I = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} d\varphi \int_0^{a \cos \varphi} f(r \cos \varphi, r \sin \varphi) r dr$$

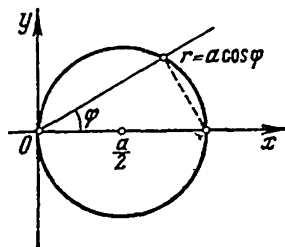


Fig. 176

(2) Let us determine the volume of the solid which is the common part of a ball of radius a and of a circular cylinder of radius $\frac{a}{2}$ on condition that the centre of the ball is on the surface of the cylinder. Here it is convenient to choose the coordinate system as shown in Fig. 177. The solid being symmetric about the planes Oxy and Oxz , it is sufficient to compute the quarter volume lying in the first octant. Thus,

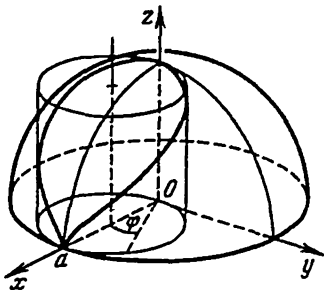


Fig. 177

$$\frac{1}{4} V = \iint_D \sqrt{a^2 - x^2 - y^2} dx dy$$

where D is the half-disc serving as half of the base of the cylinder.

In this case it is extremely convenient to pass to polar coordinates in the double integral expressing the

sought-for volume. Since the polar equation of the semicircle bounding the domain D is written $r = a \cos \varphi$ (see Example 1) we receive, integrating first with respect to r and then with respect to φ , the expression

$$\frac{1}{4} V = \int_0^{\frac{\pi}{2}} d\varphi \int_0^{a \cos \varphi} \sqrt{a^2 - r^2} r dr$$

On computing the inner integral with respect to r and integrating the result with respect to φ we obtain

$$\frac{1}{4} V = \int_0^{\frac{\pi}{2}} \left[-\frac{(a^2 - r^2)^{\frac{3}{2}}}{3} \right]_0^{a \cos \varphi} d\varphi = \frac{a^3}{3} \int_0^{\frac{\pi}{2}} (1 - \sin^3 \varphi) d\varphi = \frac{a^3}{3} \left(\frac{\pi}{2} - \frac{2}{3} \right)$$

whence

$$V = \frac{4}{3} a^3 \left(\frac{\pi}{2} - \frac{2}{3} \right)$$

III. The Euler-Poisson integral. Let us evaluate the Euler-Poisson integral

$$\int_0^{\infty} e^{-x^2} dx$$

whose convergence was established in Sec. 98.

Let D be the square with side a , and D_1 and D_2 be the quarter circles of, respectively, radii a and $a\sqrt{2}$; these domains are depicted in Fig. 178. To evaluate the desired improper integral we consider the double integrals of the auxiliary function $e^{-(x^2+y^2)}$ over the domains D , D_1 and D_2 . The integrand $e^{-(x^2+y^2)}$ being positive, the double integral of it increases as the domain of integration is extended; hence we can write the inequalities

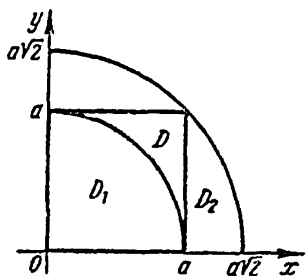


Fig. 178

$$\iint_{D_1} e^{-(x^2+y^2)} d\sigma < \iint_D e^{-(x^2+y^2)} d\sigma < \iint_{D_2} e^{-(x^2+y^2)} d\sigma \quad (*)$$

The integrals over the domains D_1 and D_2 are readily computed in polar coordinates; indeed, we have

$$\iint_{D_1} e^{-(x^2+y^2)} d\sigma = \int_0^{\frac{\pi}{2}} d\varphi \int_0^a e^{-r^2} r dr = \int_0^{\frac{\pi}{2}} \left[-\frac{e^{-r^2}}{2} \right]_0^a d\varphi = \frac{\pi}{4} (1 - e^{-a^2})$$

and for the domain D_2 we obtain $\iint_{D_2} e^{-(x^2+y^2)} d\sigma = \frac{\pi}{4} (1 - e^{-2a^2})$ by

substituting $a\sqrt{2}$ for a into the former formula.

Now we proceed to the integral over the square D :

$$\iint_D e^{-(x^2+y^2)} d\sigma = \int_0^a dx \int_0^a e^{-x^2} e^{-y^2} dy = \int_0^a e^{-x^2} dx \int_0^a e^{-y^2} dy$$

The inner integral being independent of x and of the notation of the variable of integration (this means that y can be replaced by any symbol and, in particular, by x), the right-hand side of the

last equality is nothing but the product of two equal integrals; consequently,

$$\iint_D e^{-(x^2+y^2)} d\sigma = \left(\int_0^a e^{-x^2} dx \right)^2$$

Substituting the expressions of the two integrals computed above into inequality (*) we obtain

$$\frac{\pi}{4} (1 - e^{-a^2}) < \left(\int_0^a e^{-x^2} dx \right)^2 < \frac{\pi}{4} (1 - e^{-2a^2})$$

If $a \rightarrow \infty$ the leftmost and the rightmost terms of these inequalities tend to one and the same limit $\frac{\pi}{4}$; hence, by the theorem proved in Sec. 30, we have

$$\lim_{a \rightarrow \infty} \left(\int_0^a e^{-x^2} dx \right)^2 = \left(\lim_{a \rightarrow \infty} \int_0^a e^{-x^2} dx \right)^2 = \left(\int_0^\infty e^{-x^2} dx \right)^2 = \frac{\pi}{4}$$

whence

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

The integrand being an even function, we arrive at the final result

$$\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$$

132. Applications of Double Integrals to Problems of Mechanics.

I. **The mass of a plate with variable density.** Let us consider a plate (lamina) lying in the xy -plane and occupying a domain D . The plate is supposed to be sufficiently thin so that the variation of the mass density in the direction perpendicular to the plane Oxy can be neglected.

The surface density of such a plate at a given point is the limit of the ratio of the mass of the element of area of the plate to that area as the element is contracted to this point.

The surface density thus defined depends solely on the location of the given point and thus is a function of the coordinates of the point:

$$\delta = \delta(x, y)$$

If the density were constant ($\delta = \text{const}$) the total mass of the plate would be equal to $M = \delta S$ where S is the area of the plate. If the plate is nonhomogeneous its mass is distributed with a va-

whence

$$V = \frac{4}{3} a^3 \left(\frac{\pi}{2} - \frac{2}{3} \right)$$

III. The Euler-Poisson integral. Let us evaluate the Euler-Poisson integral

$$\int_0^{\infty} e^{-x^2} dx$$

whose convergence was established in Sec. 98.

Let D be the square with side a , and D_1 and D_2 be the quarter circles of, respectively, radii a and $a\sqrt{2}$; these domains are depicted in Fig. 178. To evaluate the desired improper integral we consider the double integrals of the auxiliary function $e^{-(x^2+y^2)}$ over the domains D , D_1 and D_2 . The integrand $e^{-(x^2+y^2)}$ being positive, the double integral of it increases as the domain of integration is extended; hence we can write the inequalities

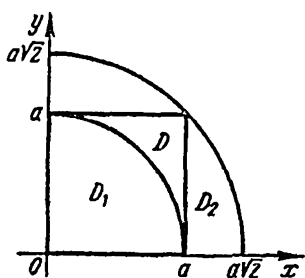


Fig. 178

$$\iint_{D_1} e^{-(x^2+y^2)} d\sigma < \iint_D e^{-(x^2+y^2)} d\sigma < \iint_{D_2} e^{-(x^2+y^2)} d\sigma \quad (*)$$

The integrals over the domains D_1 and D_2 are readily computed in polar coordinates; indeed, we have

$$\iint_{D_1} e^{-(x^2+y^2)} d\sigma = \int_0^{\frac{\pi}{2}} d\varphi \int_0^a e^{-r^2} r dr = \int_0^{\frac{\pi}{2}} \left[-\frac{e^{-r^2}}{2} \right]_0^a d\varphi = \frac{\pi}{4} (1 - e^{-a^2})$$

and for the domain D_2 we obtain $\iint_{D_2} e^{-(x^2+y^2)} d\sigma = \frac{\pi}{4} (1 - e^{-2a^2})$ by

substituting $a\sqrt{2}$ for a into the former formula.

Now we proceed to the integral over the square D :

$$\iint_D e^{-(x^2+y^2)} d\sigma = \int_0^a dx \int_0^a e^{-x^2} e^{-y^2} dy = \int_0^a e^{-x^2} dx \int_0^a e^{-y^2} dy$$

The inner integral being independent of x and of the notation of the variable of integration (this means that y can be replaced by any symbol and, in particular, by x), the right-hand side of the

If the plate is homogeneous, that is, $\delta(x, y) = \text{const}$, these formulas turn into

$$\xi = \frac{\iint_D x \, d\sigma}{S}, \quad \eta = \frac{\iint_D y \, d\sigma}{S}$$

where S is the area of the plate.

Let the reader check that if the plate is of the shape of a curvilinear trapezoid the reduction of the double integrals in the numerators to repeated integrals yields the formulas derived in Sec. 103.

III. The moments of inertia of a plate.

The moment of inertia of a material point P with mass m about an axis is the product of the mass by the square of the distance from the point P to that axis.

To find the expressions for the moments of inertia of a plate about the coordinate axes we apply the same technique as in the case of the static moments. Therefore we confine ourselves to writing down the final result for the case $\delta(x, y) \equiv 1$:

$$I_x = \iint_D y^2 \, d\sigma, \quad I_y = \iint_D x^2 \, d\sigma$$

where I_x and I_y are, respectively, the moments of inertia of the plate about Ox and Oy .

We also note that the integral $\iint_D xy \, d\sigma$ is called the *product of inertia* (of the figure D); it is denoted I_{xy} .

In mechanics we also deal with the *moment of inertia of a mass point about a point* which is equal to the product of the mass by the square of its distance from the given point. Accordingly, the *moment of inertia of the plate about the origin* is

$$I_o = \iint_D (x^2 + y^2) \, d\sigma = I_x + I_y$$

We see that this is nothing but the moment of inertia of the plate about the z -axis

§ 2. Triple Integrals

133. The Mass of a Nonhomogeneous Body. The Triple Integral. The notion of the double integral was introduced in connection with the problem of computing the volume of a curvilinear cylinder. As was shown in Sec. 132, I, the problem of finding the mass of a nonhomogeneous plate also leads to the double integral. In this section we shall state the definition of the *triple integral* proceeding from the problem of finding the mass of a non-

riable density $\delta = \delta(x, y)$. The problem is to compute the mass of the nonhomogeneous plate for a given function $\delta(x, y)$. To this end we divide the domain D occupied by the plate into subdomains $\sigma_1, \sigma_2, \dots, \sigma_n$ with areas $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ (Fig. 179), choose a point $P_i(x_i, y_i)$ within each subdomain and consider the density of mass distribution at all the points of the subdomain

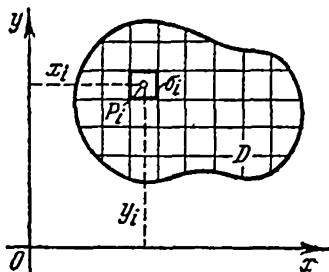


Fig. 179

σ_i ($i = 1, 2, \dots, n$) as being approximately constant and equal to the value of the density $\delta(x_i, y_i)$ at the chosen point. Then an approximation to the total mass of the plate can be written as the integral sum

$$M_n = \sum_{i=1}^n \delta(x_i, y_i) \Delta\sigma_i \quad (*)$$

To determine the exact value of the mass we should compute the limit of sum (*) as $n \rightarrow \infty$ on condition that every subdomain is contracted to a point. This results in

$$M = \iint_D \delta(x, y) d\sigma$$

II. The static moments and the coordinates of the centre of gravity of a plate. We now proceed to compute the static moments of a given plate about the coordinate axes. To this end, we consider the masses of the subdomains σ_i ($i = 1, 2, \dots, n$) as concentrated at the points $P_i(x_i, y_i)$ and determine the static moments of the resulting system of mass points (see Sec. 103):

$$M_x^{(n)} = \sum_{i=1}^n y_i \delta(x_i, y_i) \Delta\sigma_i, \quad M_y^{(n)} = \sum_{i=1}^n x_i \delta(x_i, y_i) \Delta\sigma_i$$

As usual, on passing to the limit and replacing the integral sums by the corresponding integrals we obtain the static moments of the plate:

$$M_x = \iint_D y \delta(x, y) d\sigma, \quad M_y = \iint_D x \delta(x, y) d\sigma$$

Taking advantage of the results of Sec. 103 we now find the coordinates of the centre of gravity of the plate:

$$\xi = \frac{M_y}{M} = \frac{\iint_D x \delta(x, y) d\sigma}{\iint_D \delta(x, y) d\sigma}, \quad \eta = \frac{M_x}{M} = \frac{\iint_D y \delta(x, y) d\sigma}{\iint_D \delta(x, y) d\sigma}$$

If the plate is homogeneous, that is, $\delta(x, y) = \text{const}$, these formulas turn into

$$\xi = \frac{\iint_D x \, d\sigma}{S}, \quad \eta = \frac{\iint_D y \, d\sigma}{S}$$

where S is the area of the plate.

Let the reader check that if the plate is of the shape of a curvilinear trapezoid the reduction of the double integrals in the numerators to repeated integrals yields the formulas derived in Sec. 103.

III. The moments of inertia of a plate.

The moment of inertia of a material point P with mass m about an axis is the product of the mass by the square of the distance from the point P to that axis.

To find the expressions for the moments of inertia of a plate about the coordinate axes we apply the same technique as in the case of the static moments. Therefore we confine ourselves to writing down the final result for the case $\delta(x, y) \equiv 1$:

$$I_x = \iint_D y^2 \, d\sigma, \quad I_y = \iint_D x^2 \, d\sigma$$

where I_x and I_y are, respectively, the moments of inertia of the plate about Ox and Oy .

We also note that the integral $\iint_D xy \, d\sigma$ is called the *product of inertia* (of the figure D); it is denoted I_{xy} .

In mechanics we also deal with the *moment of inertia of a mass point about a point* which is equal to the product of the mass by the square of its distance from the given point. Accordingly, the *moment of inertia of the plate about the origin* is

$$I_0 = \iint_D (x^2 + y^2) \, d\sigma = I_x + I_y$$

We see that this is nothing but the moment of inertia of the plate about the z -axis

§ 2. Triple Integrals

133. The Mass of a Nonhomogeneous Body. The Triple Integral. The notion of the double integral was introduced in connection with the problem of computing the volume of a curvilinear cylinder. As was shown in Sec. 132, I, the problem of finding the mass of a nonhomogeneous plate also leads to the double integral. In this section we shall state the definition of the *triple integral* proceeding from the problem of finding the *mass of a non-*

homogeneous physical body.¹ Since there is a complete analogy between the definitions of the double and the triple integrals we shall limit ourselves to a brief account of the theory of the triple integral. Let us consider a body occupying a spatial domain Ω ; we shall suppose that the density of mass distribution within the body is a known function continuous throughout Ω :

$$\delta = \delta(x, y, z)$$

The unit of density is kg/m^3 .

Let us break up the domain Ω in an arbitrary fashion into n parts and denote by $\Delta v_1, \Delta v_2, \dots, \Delta v_n$ the volumes of these parts. Choosing a point $P_l(x_l, y_l, z_l)$ in each subdomain and assuming that the density is approximately constant and equal to its value at the point P_l within every subdomain we can write down the approximation

$$M_n = \sum_{l=1}^n \delta(x_l, y_l, z_l) \Delta v_l \quad (*)$$

to the mass of the body. The limit of sum (*) as $n \rightarrow \infty$, on condition that each subdomain is contracted to a point (that is, as the *diameters*² of the subdomains tend to zero) is equal to the total mass M of the body:

$$M = \lim_{n \rightarrow \infty} \sum_{l=1}^n \delta(x_l, y_l, z_l) \Delta v_l = \iiint_{\Omega} \delta(x, y, z) dv$$

The sum (*) is called an *integral sum* and its limit is called the *triple integral of the function $\delta(x, y, z)$ over the spatial domain Ω* .

There are various other problems leading to the computation of a triple integral. Therefore we shall consider the general form of the triple integral

$$\iiint_{\Omega} f(x, y, z) dv = \lim_{n \rightarrow \infty} \sum_{l=1}^n f(x_l, y_l, z_l) \Delta v_l$$

where $f(x, y, z)$ is an arbitrary continuous function in a finite domain Ω of volume V ; such a domain is usually bounded by one or several closed surfaces.

¹ As was said in Sec. 107, it is impossible to give a visual *geometrical* interpretation of a function of three independent variables. That is why we introduce the notion of the triple integral, without geometrical considerations, on the basis of the mechanical problem of determining the mass of a physical body.

² The *diameter* of a spatial figure is defined by analogy with the diameter of a plane figure (see p. 438). If the diameter of a spatial subdomain tends to zero its volume also tends to zero.

In this integral with respect to z the variables x and y are regarded as being constant in the integration process.

The value of the sought-for triple integral can now be found if we take the integral of the function $F(x, y)$ on condition that the point $P(x, y)$ runs through the domain D , that is, if we compute the double integral

$$\iint_D F(x, y) dx dy$$

Thus, the triple integral I can be represented in the form

$$I = \iint_D dx dy \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz$$

Furthermore, reducing the double integral over D to an iterated integral with inner integration with respect to y and outer integration with respect to x we obtain

$$I = \int_a^b dx \int_{y_1(x)}^{y_2(x)} dy \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz \quad (*)$$

where $y_1(x)$ and $y_2(x)$ are the ordinates of the points of intersection of the straight line $x = \text{const}$ in the xy -plane with the boundary of the domain D and a and b are the abscissas of the end points of the interval of the x -axis onto which the domain D is projected.

We see that the computation of the triple integral over the domain Ω reduces to three consecutive definite integrations.

Formula (*) also remains valid for domains of the form of a cylindrical surface with elements parallel to Oz and by two surfaces $z = z_1(x, y)$ and $z = z_2(x, y)$ (see Fig. 181).

If the domain of integration is the interior of a parallelepiped with faces parallel to the coordinate planes (Fig. 182) the limits of integration are constant in all the three integrals:

$$\iiint_{\Omega} f(x, y, z) dx dy dz = \int_a^b dx \int_c^d dy \int_k^l f(x, y, z) dz$$

In this case the order of integration can be changed in an arbitrary

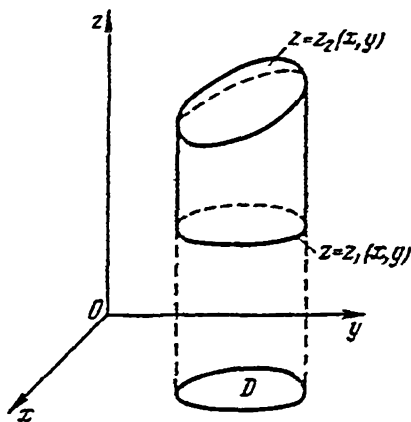


Fig. 181

to the product of the differentials of the variables of integration:¹

$$dv = dx dy dz$$

Accordingly, we shall write

$$I = \iiint_{\Omega} f(x, y, z) dx dy dz$$

Now we proceed to establish a rule for computing such an integral. We shall assume that the domain of integration Ω is of the shape depicted in Fig. 180.²

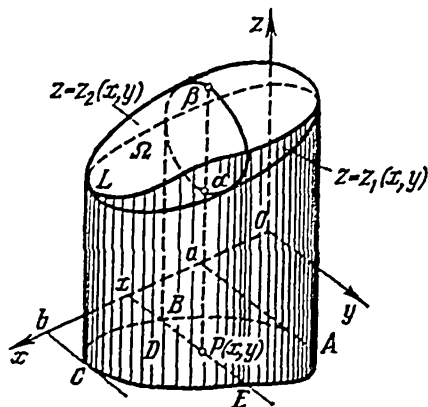


Fig. 180

as the orthogonal projection of the spatial domain Ω on that plane; the line L is then projected on the boundary of the domain D .

Let us begin with the integration in the direction of Oz . To this end, we integrate the function $f(x, y, z)$ along a line segment enclosed in Ω and parallel to Oz , its extension meeting the domain D at a point $P(x, y)$ (see the line segment $\alpha\beta$ in Fig. 180). For given values of x and y the variable z ranges from $z_1(x, y)$ (the z -coordinate of the point α) to $z_2(x, y)$ (the z -coordinate of the point β).

The result of this first integration is an expression dependent on the point $P(x, y)$; we denote it by $F(x, y)$:

$$F(x, y) = \int_{z_1(x, y)}^{z_2(x, y)} f(x, y, z) dz$$

¹ In connection with this assertion see the footnote on p. 442 which should be restated for the case of a spatial domain.

² If the domain of integration Ω is of a more complicated form it should be divided into a number of parts of the indicated shape (provided it is possible) and the given integral should be computed as the sum of the integrals taken over these parts.

and $x + y = 1$:

$$\begin{aligned} I &= \int_0^1 dx \int_0^{1-x} \left[(x+y) - (x+y)^2 + \frac{(1-x-y)^2}{2} \right] dy = \\ &= \int_0^1 dx \left[\frac{(x+y)^2}{2} - \frac{(x+y)^3}{3} - \frac{(1-x-y)^3}{6} \right] \Big|_0^{1-x} = \\ &= \int_0^1 \left[\frac{1}{2} - \frac{x^2}{2} - \frac{1}{3} + \frac{x^3}{3} + \frac{(1-x)^3}{6} \right] dx = \frac{1}{8} \end{aligned}$$

The rule for change of variables in a triple integral is derived on the basis of the same principles as in the case of the double integral. Here we do not present the general formula and confine ourselves to considering the cases of the transformation to cylindrical and spherical coordinates which admit of a visual geometrical interpretation.

II. Cylindrical coordinates.

Let us consider a domain Ω in space where *cylindrical coordinates* (r, φ, z) are introduced. In these coordinates the position of the moving point M in space is determined by the polar coordinates (r, φ) of its projection P on the plane Oxy and by its z -coordinate. Choosing the coordinate axes as shown in Fig. 184 we can write down the relations connecting the Cartesian and cylindrical coordinates of the point M :

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z \quad (*)$$

Let us divide the domain Ω into subdomains v_i with the aid of three systems of coordinate surfaces $r = \text{const}$, $\varphi = \text{const}$ and $z = \text{const}$ which are, respectively, circular cylindrical surfaces with Oz as axis, half-planes passing through the axis Oz and planes parallel to the plane Oxy . The subdomains v_i are right cylinders of the form of MN (Fig. 184). The volume of such a cylinder being equal to the product of the base area by the altitude, we obtain the following expression for the element of volume in cylindrical coordinates:

$$dv = r dr d\varphi dz$$

The transformation of the triple integral $\iiint_{\Omega} f(x, y, z) dv$ to cylin-

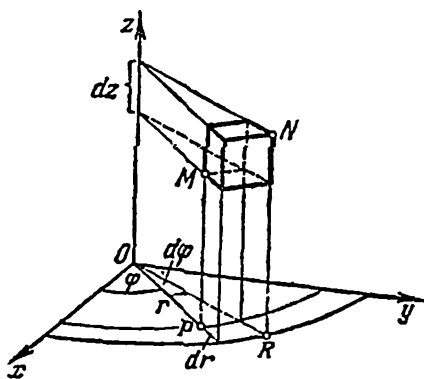


Fig. 184

trary manner with the retention of the limits of integration for each variable.

In the general case a change of the order of integration in the above-described procedure (for instance, the passage to the case

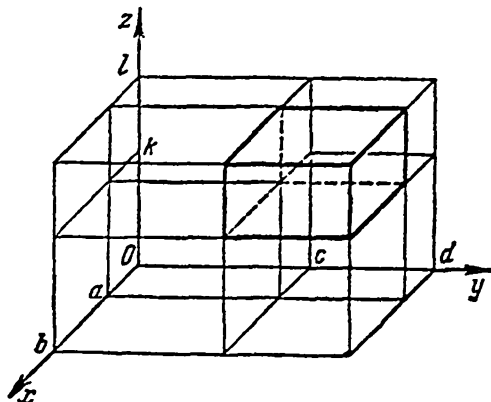


Fig. 182

when we first integrate with respect to y and then with respect to x and z in the plane Oxz) results in a change of the order of integration in the threefold iterated integral and in a change of the limits of integration for each variable.

Example. Let us evaluate the triple integral

$$I = \iiint_{\Omega} (x+y+z) dx dy dz$$

over the domain Ω bounded by the coordinate planes

$$x=0, \quad y=0, \quad z=0$$

and the plane $x+y+z=1$ (this is the pyramid depicted in Fig. 183).

The integration with respect to z goes from $z=0$ to $z=1-x-y$. Therefore, denoting the projection of the domain Ω on the xy -plane by D , we obtain

$$\begin{aligned} I &= \iint_D dx dy \int_0^{1-x-y} (x+y+z) dz = \iint_D \left[(x+y)z + \frac{z^2}{2} \right]_0^{1-x-y} dx dy = \\ &= \iint_D \left[(x+y) - (x+y)^2 + \frac{(1-x-y)^2}{2} \right] dx dy \end{aligned}$$

Now we must set up the limits of integration for the domain D which is the triangle with sides having the equations $x=0$, $y=0$

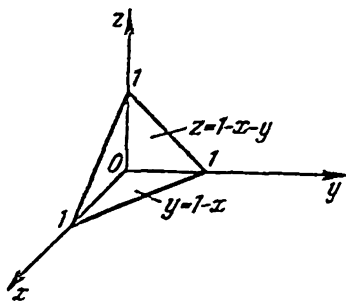


Fig. 183

and $x + y = 1$:

$$\begin{aligned}
 I &= \int_0^1 dx \int_0^{1-x} \left[(x+y) - (x+y)^2 + \frac{(1-x-y)^2}{2} \right] dy = \\
 &= \int_0^1 dx \left[\frac{(x+y)^2}{2} - \frac{(x+y)^3}{3} - \frac{(1-x-y)^3}{6} \right] \Big|_0^{1-x} = \\
 &= \int_0^1 \left[\frac{1}{2} - \frac{x^2}{2} - \frac{1}{3} + \frac{x^3}{3} + \frac{(1-x)^3}{6} \right] dx = \frac{1}{8}
 \end{aligned}$$

The rule for change of variables in a triple integral is derived on the basis of the same principles as in the case of the double integral. Here we do not present the general formula and confine ourselves to considering the cases of the transformation to cylindrical and spherical coordinates which admit of a visual geometrical interpretation.

II. Cylindrical coordinates.

Let us consider a domain Ω in space where *cylindrical coordinates* (r, φ, z) are introduced. In these coordinates the position of the moving point M in space is determined by the polar coordinates (r, φ) of its projection P on the plane Oxy and by its z -coordinate. Choosing the coordinate axes as shown in Fig. 184 we can write down the relations connecting the Cartesian and cylindrical coordinates of the point M :

$$x = r \cos \varphi, \quad y = r \sin \varphi, \quad z = z \quad (*)$$

Let us divide the domain Ω into subdomains v_i with the aid of three systems of *coordinate surfaces* $r = \text{const}$, $\varphi = \text{const}$ and $z = \text{const}$ which are, respectively, circular cylindrical surfaces with Oz as axis, half-planes passing through the axis Oz and planes parallel to the plane Oxy . The subdomains v_i are right cylinders of the form of MN (Fig. 184). The volume of such a cylinder being equal to the product of the base area by the altitude, we obtain the following expression for the element of volume in cylindrical coordinates:

$$dv = r dr d\varphi dz$$

The transformation of the triple integral $\iiint_{\Omega} f(x, y, z) dv$ to cylin-

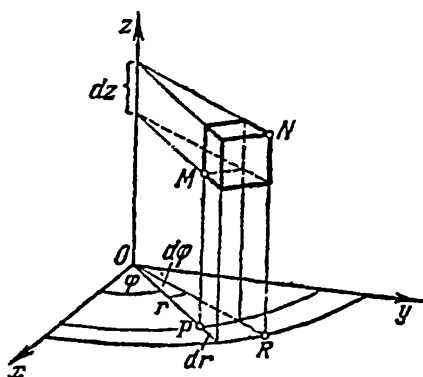


Fig. 184

drical coordinates is performed by analogy with the transformation of the double integral from Cartesian to polar coordinates. To this end, the variables x, y, z in the function $f(x, y, z)$ are changed according to formulas (*) and the element of volume dv is replaced by $r dr d\varphi dz$; in the new variables r, φ, z the triple integral is taken over the corresponding domain Ω^* in the auxiliary Cartesian coordinates $O_1 r \varphi z$. The final result is written as

$$\iiint_{\Omega} f(x, y, z) dv = \iiint_{\Omega^*} f(r \cos \varphi, r \sin \varphi, z) r dr d\varphi dz$$

In many cases encountered in practice there is no need in constructing the domain Ω^* since the limits of integration in the new coordinates can be usually set up in accordance with the shape of the original domain Ω . As a rule, the inner integration is performed with respect to the variable z ; the equations of the surfaces bounding the domain Ω should be written in the cylindrical coordinates. In particular, if the domain of integration is the interior of a right cylinder $r \leq R, 0 \leq z \leq h$, all the limits of integration are constant:

$$I = \int_0^{2\pi} d\varphi \int_0^R r dr \int_0^h f(r \cos \varphi, r \sin \varphi, z) dz$$

If, in the case of such a domain, the order of integration is changed the limits of integration for each of the arguments remain invariable.

Example. Let us evaluate the integral $\iiint_{\Omega} z dv$ over the domain Ω bounded

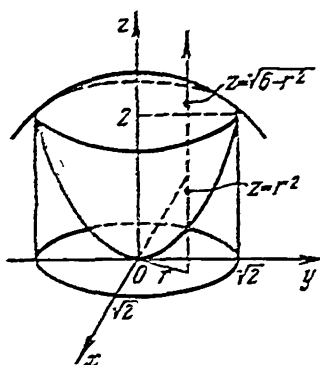


Fig. 185

below by the paraboloid $z = x^2 + y^2$ and bounded above by the sphere $x^2 + y^2 + z^2 = 6$ (see Fig. 185). In the cylindrical coordinates the equations of these surfaces are, respectively, $z = r^2$ and $r^2 + z^2 = 6$. Its line of intersection is a circle in the plane $z = 2$; the radius of this circle is equal to $\sqrt{2}$ (these values appear when the system of equations $z = r^2, r^2 + z^2 = 6$ is solved). Consequently,

$$\begin{aligned} \iiint_{\Omega} z dv &= \int_0^{2\pi} d\varphi \int_0^{\sqrt{2}} r dr \int_{r^2}^{\sqrt{6-r^2}} z dz = \\ &= \int_0^{2\pi} d\varphi \int_0^{\sqrt{2}} \frac{z^2}{2} \Big|_{r^2}^{\sqrt{6-r^2}} r dr = \pi \int_0^{\sqrt{2}} (6 - r^2 - r^4) r dr = \frac{11}{3} \pi \end{aligned}$$

III. Spherical coordinates. Now let us perform the transformation from Cartesian coordinates (x, y, z) to spherical coordinates

(r, θ, φ) in the triple integral $\iiint_{\Omega} f(x, y, z) dv$ over a spatial domain Ω . In spherical coordinates the position of the variable point M is specified by its distance r from the origin (i.e. by the length of the radius vector of the point), the angle θ between the radius vector of the point and the z -axis and by the angle φ between the projection of the radius vector on the plane Oxy and the x -axis (see Fig. 186). The angle θ varies from 0 to π and φ from 0 to 2π .¹

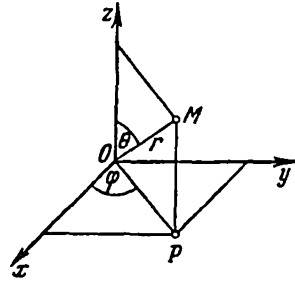


Fig. 186

The connection between the Cartesian and spherical coordinates is readily established. Fig. 186 indicates that

$$MP = r \sin \left(\frac{\pi}{2} - \theta \right) = r \cos \theta, \quad OP = r \cos \left(\frac{\pi}{2} - \theta \right) = r \sin \theta$$

$$x = OP \cos \varphi \quad \text{and} \quad y = OP \sin \varphi$$

It follows that

$$x = r \sin \theta \cos \varphi, \quad y = r \sin \theta \sin \varphi, \quad z = r \cos \theta \quad (**)$$

Let us break up the domain Ω into subdomains v_i by means of three systems of coordinate surfaces $r = \text{const}$, $\theta = \text{const}$ and $\varphi = \text{const}$. These surfaces are, respectively, concentric spheres with centre at the origin, half-planes passing through the z -axis and circular cones with common vertex at the origin and the positive or negative half of Oz as axis. The subdomains v_i are of the shape of a curvilinear "hexahedron" MN shown in heavy line in Fig. 187. Discarding the infinitesimals of higher order we can approximately consider the hexahedron MN as a rectangular parallelepiped with dimensions dr in the direction of the

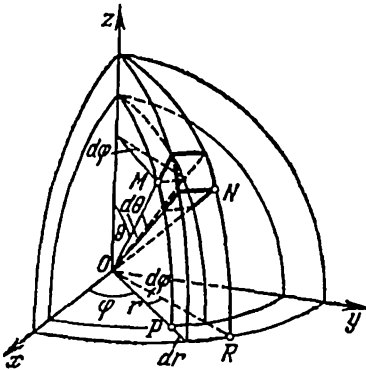


Fig. 187

radius vector, $r d\theta$ in the meridional direction and $r \sin \theta d\varphi$ in the direction of the parallel. Then for the element of volume in spheri-

¹ In cartography the position of the point on the surface of the Earth is specified by the angle φ (called the longitude of the point) and by the angle $\theta' = \frac{\pi}{2} - \theta$ (the latitude of the point). The longitude then varies from 0 to 2π

(or from $-\pi$ to π) and the latitude from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$.

cal coordinates we obtain the expression

$$dv = r^2 \sin \theta \, dr \, d\varphi \, d\theta$$

Changing the variables x, y, z in the triple integral according to formulas (***) and taking the above expression of the element of volume dv we pass to the integration over the corresponding domain Ω^* in the auxiliary Cartesian coordinates $O_1 r \theta \varphi$:

$$\begin{aligned} \iiint_{\Omega} f(x, y, z) \, dv &= \\ &= \iiint_{\Omega^*} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta \, dr \, d\varphi \, d\theta \end{aligned}$$

The application of spherical coordinates becomes particularly convenient when the domain of integration Ω is a ball with centre at the origin or a spherical shell bounded by two concentric spheres of radii R_1 and R_2 ($R_1 < R_2$). In the latter case the limits of integration are set up as

$$\int_0^{\pi} \sin \theta \, d\theta \int_0^{2\pi} d\varphi \int_{R_1}^{R_2} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \, dr$$

If Ω is a ball of radius R we should put $R_2 = R$ and $R_1 = 0$ in the latter formula. The limits of integration we have spoken about have been set up in accordance with the shape of the domain Ω without resorting to the auxiliary Cartesian coordinates $O_1 r \theta \varphi$.

Example. Let us compute the triple integral $I = \iiint_{\Omega} xyz \, dv$ where Ω is the part of the ball $x^2 + y^2 + z^2 \leq R^2$ in the first coordinate trihedral. Setting up the limits of integration we find

$$I = \int_0^{\frac{\pi}{2}} \sin^3 \theta \cos \theta \, d\theta \int_0^{\frac{\pi}{2}} \sin \varphi \cos \varphi \, d\varphi \int_0^R r^3 \, dr = \frac{1}{4} \cdot \frac{1}{2} \cdot \frac{R^4}{6} = \frac{R^4}{48}$$

In conclusion we note that an appropriate choice of the coordinate system depends both on the shape of the domain of integration and on the properties of the integrand, and it is impossible to elaborate a general rule for such a choice. It is sometimes advisable to represent the given integral in different coordinates and then decide which of them is more suitable.

135. Applications of Triple Integrals. The application of triple integrals to computing static moments and moments of inertia of spatial bodies is based on the same principles as the application of double integrals to the moments of plane plates (see Sec. 132). To determine the coordinates of the centre of gravity of a body it is necessary to find the static moments

with respect to the coordinate planes Oxy , Oxz and Oyz (we denote them M_{xy} , M_{xz} and M_{yz}). Repeating the argument of Sec. 132, II (let the reader consider this question in detail) we derive the following formulas for the coordinates ξ , η , ζ of the centre of gravity of a nonhomogeneous physical body occupying a domain Ω and having a (volume) density of mass distribution $\delta(x, y, z)$:

$$\xi = \frac{M_{yz}}{M} = \frac{\iiint_{\Omega} x \delta \, dv}{\iiint_{\Omega} \delta \, dv}, \quad \eta = \frac{M_{xz}}{M} = \frac{\iiint_{\Omega} y \delta \, dv}{\iiint_{\Omega} \delta \, dv}, \quad \zeta = \frac{M_{xy}}{M} = \frac{\iiint_{\Omega} z \delta \, dv}{\iiint_{\Omega} \delta \, dv}$$

If the body is homogeneous, i.e. $\delta = \text{const}$, these formulas are simplified:

$$\xi = \frac{\iiint_{\Omega} x \, dv}{V}, \quad \eta = \frac{\iiint_{\Omega} y \, dv}{V}, \quad \zeta = \frac{\iiint_{\Omega} z \, dv}{V}$$

where V is the volume of the body.

Example. Let us determine the coordinates of the centre of gravity of the homogeneous half-ball Ω specified by the conditions

$$x^2 + y^2 + z^2 \leq R^2, \quad z \geq 0$$

The coordinates ξ and η of the centre of gravity are equal to zero since the half-ball is symmetric with respect to the axis Oz (it is a solid of revolution about Oz).

The integral $\iiint_{\Omega} z \, dv$ is readily computed in spherical coordinates:

$$\iiint_{\Omega} z \, dv = \int_0^{\frac{\pi}{2}} \sin \theta \cos \theta \, d\theta \int_0^{2\pi} d\varphi \int_0^R r^3 \, dr = \frac{1}{2} 2\pi \frac{R^4}{4} = \frac{1}{4} \pi R^4$$

The volume of the half-ball being equal to $\frac{2}{3} \pi R^3$, we get $\zeta =$

$$= \frac{\frac{1}{4} \pi R^4}{\frac{2}{3} \pi R^3} = \frac{3}{8} R.$$

Let us proceed to find the moments of inertia of a body about the coordinate axes (see Sec. 132, III). Since the squares of the distances from the moving point $P(x, y, z)$ to the axes Ox , Oy and Oz are, respectively, $y^2 + z^2$, $x^2 + z^2$ and $x^2 + y^2$, we obtain (for the sake of simplicity we put $\delta \equiv 1$) the formulas

$$I_x = \iiint_{\Omega} (y^2 + z^2) \, dv, \quad I_y = \iiint_{\Omega} (x^2 + z^2) \, dv, \quad I_z = \iiint_{\Omega} (x^2 + y^2) \, dv$$

As in the case of a plane region, the integrals

$$I_{xy} = \iiint_{\Omega} xy \, dv, \quad I_{yz} = \iiint_{\Omega} yz \, dv, \quad I_{zx} = \iiint_{\Omega} zx \, dv$$

are termed the *products of inertia*.

For the moment of inertia about the origin we derive the expression

$$I_o = \iiint_{\Omega} (x^2 + y^2 + z^2) \, dv$$

If the body is *nonhomogeneous* each of these formulas involves the additional factor $\delta(x, y, z)$ (the density of the body at the point $P(x, y, z)$) under the integral sign.

Example. Let us compute the moment of inertia of a homogeneous ball of radius R with respect to its centre. In this case it is convenient to use spherical coordinates with origin at the centre of the ball. This yields

$$I_o = \delta \int_0^{\pi} \sin \theta \, d\theta \int_0^{2\pi} d\varphi \int_0^R r^2 r^2 \, dr = \frac{4\pi\delta R^5}{5} = \frac{3}{5} MR^2$$

where M is the mass of the ball.

Since the moments of inertia of the ball relative to the coordinate axes (passing through its centre) are obviously equal to each other, we can make use of the relation $I_x + I_y + I_z = 2I_o$ to receive

$$I_x = I_y = I_z = \frac{2}{5} MR^2$$

The moments of inertia of a physical body play an important role in the computation of the kinetic energy of a rotary motion of the body about the corresponding axis. Let a body Ω rotate about the axis Oz with constant angular velocity ω . We shall find the kinetic energy J_z of the body. As is known, the kinetic energy of a mass point is equal to $\frac{1}{2}mv^2$ where m is the mass of the point and v the magnitude of its velocity. The kinetic energy of a system of material points is defined as the sum of the kinetic energies of the constituent mass points, and the kinetic energy of a body as the sum of the kinetic energies of the parts the whole body is divided into. This makes it possible to apply the integral to computing the kinetic energy.

Let $P(x, y, z)$ be a variable point of the domain Ω occupied by the body and dv the volume of an infinitesimal neighbourhood of this point. The magnitude of the linear velocity v of the point P in the rotary motion about the z -axis is equal to $\omega\sqrt{x^2 + y^2}$, and

therefore the kinetic energy of the element dv of the body Ω is expressed as

$$\frac{1}{2} \delta(P) dv \omega^2 (x^2 + y^2)$$

where $\delta(P) = \delta(x, y, z)$ is the mass density of the body at the point P . Consequently, for the kinetic energy of the whole body we obtain the expression

$$J_z = \iiint_{\Omega} \frac{1}{2} \omega^2 (x^2 + y^2) \delta(P) dv = \frac{1}{2} \omega^2 \iiint_{\Omega} (x^2 + y^2) \delta(P) dv$$

that is,

$$J_z = \frac{1}{2} \omega^2 I_z$$

Thus, the kinetic energy of a body rotating about an axis with constant angular velocity is equal to half the product of the square of the angular velocity by the moment of inertia of the body with respect to the axis of revolution.

§ 3*. Integrals Dependent on Parameters

136*. Integrals with Finite Limits. We have already dealt with integrals involving integrands dependent on two variables and integration with respect to one of the variables under the assumption that the other variable is considered constant in the integration process. We had such a situation in Sec. 115 when a function was reconstructed from its total differential and in Sec. 130 when the double integral was evaluated by means of successive definite integrations. Since integrals of this kind are widely encountered in mathematical analysis and other divisions of higher mathematics (for instance, in the theory of equations of mathematical physics) we shall consider them in more detail in this section.

Let there be given a function $f(x, \lambda)$ defined and continuous for all the values of x and λ satisfying the conditions

$$a \leq x \leq b \quad \text{and} \quad \lambda_1 \leq \lambda \leq \lambda_2$$

(These conditions specify a rectangle in the plane $Ox\lambda$.)

If the integral $\int_a^b f(x, \lambda) dx$ has been computed for every value $\lambda \in [\lambda_1, \lambda_2]$ then to each λ there corresponds a definite value of the integral which thus is a function of the parameter λ ; denoting this function $F(\lambda)$ we write

$$F(\lambda) = \int_a^b f(x, \lambda) dx \quad (*)$$

In this section we suppose that the interval of integration $[a, b]$ is *finite* (the case of an infinite interval of integration will be treated in the next section). The interval $[\lambda_1, \lambda_2]$ serving as the domain of definition of the function $F(\lambda)$ can be finite or infinite.

Examples. (1) If $\lambda \neq 0$ we have $\int_0^1 e^{\lambda x} dx = \frac{e^{\lambda x}}{\lambda} \Big|_0^1 = \frac{e^\lambda - 1}{\lambda}$. For $\lambda = 0$ this integral is equal to 1. Here the function $F(\lambda)$ is defined and continuous for all the values of λ (let the reader check the continuity for $\lambda = 0$).

(2) Considering the values of λ satisfying the condition $|\lambda| > 1$ we can write $\int_0^1 \frac{dx}{\sqrt{\lambda^2 - x^2}} = \arcsin \frac{x}{\lambda} \Big|_0^1 = \arcsin \frac{1}{\lambda}$. For $\lambda = \pm 1$ the integral becomes improper (in this case the integrand has an infinite discontinuity at the point $x = 1$) but convergent.

(3) For $\lambda \in [-1, 1]$ we have $\int_1^2 \frac{dx}{\sqrt{x^2 - \lambda^2}} = \ln(x + \sqrt{x^2 - \lambda^2}) \Big|_1^2 = \ln \frac{2 + \sqrt{4 - \lambda^2}}{1 + \sqrt{1 - \lambda^2}}$.

It should be stressed once again that if the function $f(x, \lambda)$ and the interval $[a, b]$ are fixed integral (*) depends solely on λ . Our further aim is to investigate the properties of the function $F(\lambda)$ termed an *integral dependent on the parameter λ* . First of all we prove the following fundamental theorem:

Theorem 1. If $f(x, \lambda)$ is a continuous function for $a \leq x \leq b$, $\lambda_1 \leq \lambda \leq \lambda_2$ the function $F(\lambda) = \int_a^b f(x, \lambda) dx$ is continuous in the interval $[\lambda_1, \lambda_2]$.

Proof. To simplify the proof we shall suppose additionally that the function $f(x, \lambda)$ possesses the partial derivative $f'_\lambda(x, \lambda)$ bounded for all the values of x and λ in question: $|f'_\lambda(x, \lambda)| \leq M$ (this additional condition is not mentioned in the statement of the theorem which also remains true when it is violated).

Let λ and $\lambda + \Delta\lambda$ be two values of the parameter belonging to the interval $[\lambda_1, \lambda_2]$. By definition, we have

$$F(\lambda) = \int_a^b f(x, \lambda) dx \text{ and } F(\lambda + \Delta\lambda) = \int_a^b f(x, \lambda + \Delta\lambda) dx$$

The increment of the function $F(\lambda)$ is equal to

$$F(\lambda + \Delta\lambda) - F(\lambda) = \int_a^b [f(x, \lambda + \Delta\lambda) - f(x, \lambda)] dx \quad (**)$$

Let us rewrite the expression of the integrand in (**) by using Lagrange's formula of finite increments:

$$f(x, \lambda + \Delta\lambda) - f(x, \lambda) = f'_\lambda(x, \xi) \Delta\lambda$$

where $\xi \in (\lambda, \lambda + \Delta\lambda)$. The derivative f'_λ being bounded, we arrive at the inequality

$$|f(x, \lambda + \Delta\lambda) - f(x, \lambda)| \leq M |\Delta\lambda|$$

Since the absolute value of an integral does not exceed the integral of the modulus of the integrand (see Sec. 90, II), we have

$$\begin{aligned} |F(\lambda + \Delta\lambda) - F(\lambda)| &= \left| \int_a^b [f(x, \lambda + \Delta\lambda) - f(x, \lambda)] dx \right| \leq \\ &\leq \int_a^b M |\Delta\lambda| dx = M(b-a) |\Delta\lambda| \end{aligned}$$

It follows that $F(\lambda + \Delta\lambda) - F(\lambda) \rightarrow 0$ as $\Delta\lambda \rightarrow 0$, that is, the function $F(\lambda)$ is continuous.

In what follows we shall use this theorem in its original formulation (without stipulating the existence and the boundedness of the derivative f'_λ , which is inessential although this hypothesis was taken into account in the present proof). It should also be noted that the condition that the interval $[a, b]$ is *finite* was essentially used in this proof.

The condition expressing the continuity of the integral $\int_a^b f(x, \lambda) dx$ dependent on the parameter λ can be written in the form

$$\lim_{\lambda \rightarrow \lambda_0} \int_a^b f(x, \lambda) dx = \int_a^b f(x, \lambda_0) dx \quad (***)$$

indicating that the limit of the continuous function $F(\lambda)$ as $\lambda \rightarrow \lambda_0$ is equal to the particular value of this function for $\lambda = \lambda_0$.

When using this theorem we should check that all its conditions are fulfilled since if otherwise our conclusions may be incorrect. For instance, let

$$F(\lambda) = \int_0^1 \frac{\lambda dx}{x^2 + \lambda^2} = \arctan \frac{x}{\lambda} \Big|_0^1 = \arctan \frac{1}{\lambda}$$

This function is continuous for all the values of the parameter λ except for $\lambda = 0$; the matter is that the integrand $f(x, \lambda) = \frac{\lambda}{x^2 + \lambda^2}$ is a discontinuous function for $\lambda = 0$ and $x = 0$.

For the function $F(\lambda)$ to be differentiable some additional requirements should be imposed on the function $f(x, \lambda)$:

Theorem II. If $f(x, \lambda)$ is a function defined and continuous for $a \leq x \leq b$, $\lambda_1 \leq \lambda \leq \lambda_2$ together with its partial derivative $f'_\lambda(x, \lambda)$ the function $F(\lambda)$ possesses a continuous derivative with respect to λ expressed by the formula

$$F'(\lambda) = \frac{d}{d\lambda} \int_a^b f(x, \lambda) dx = \int_a^b f'_\lambda(x, \lambda) dx$$

Proof. We shall again take advantage of formula (**) expressing the increment of the function $F(\lambda)$. On dividing its both sides by $\Delta\lambda$ and transforming the integrand according to Lagrange's theorem on finite increments we receive

$$\frac{F(\lambda + \Delta\lambda) - F(\lambda)}{\Delta\lambda} = \int_a^b f'_\lambda(x, \xi) dx$$

Since $\xi \in (\lambda, \lambda + \Delta\lambda)$, we have $\xi \rightarrow \lambda$ as $\Delta\lambda \rightarrow 0$. Therefore, passing to the limit as $\Delta\lambda \rightarrow 0$, we obtain

$$\lim_{\Delta\lambda \rightarrow 0} \frac{F(\lambda + \Delta\lambda) - F(\lambda)}{\Delta\lambda} = \lim_{\xi \rightarrow \lambda} \int_a^b f'_\lambda(x, \xi) dx$$

By the hypothesis, the derivative $f'_\lambda(x, \lambda)$ is *continuous*; consequently, the limit on the right-hand side of the latter relation exists (see Formula (***)). Hence, the limit on the left-hand side also exists, these limits being equal to each other and expressing the derivative $F'(\lambda)$:

$$F'(\lambda) = \int_a^b f'_\lambda(x, \lambda) dx$$

The theorem has been proved. It can be stated briefly as

the derivative of an integral with respect to a parameter it depends on is equal to the integral of the derivative of the integrand with respect to that parameter. This proposition is known as *Leibniz' rule*; it says that, under certain conditions, it is allowable to interchange the operation of integration with respect to x and the operation of differentiation with respect to λ . This rule was already mentioned in Sec. 115 in connection with the reconstruction of a function from its total differential.

Leibniz' rule sometimes facilitates the computation of complicated definite integrals.

Example. We have $\int_0^1 \frac{dx}{x^2+a^2} = \frac{1}{a} \arctan \frac{x}{a} \Big|_0^1 = \frac{1}{a} \arctan \frac{1}{a}$. Differen-

tiating both sides of this equality with respect to the parameter a and applying Leibniz' rule to the differentiation of the integral (this means that the integral is *differentiated under the sign of integration*, that is, the integrand itself is differentiated with respect to a) we obtain

$$\int_0^1 \frac{-2a dx}{(x^2+a^2)^2} = -\frac{1}{a^2} \arctan \frac{1}{a} + \frac{1}{a} \frac{1}{1+\left(\frac{1}{a}\right)^2} \left(-\frac{1}{a^2}\right)$$

whence

$$\int_0^1 \frac{dx}{(x^2+a^2)^2} = \frac{1}{2a^3} \arctan \frac{1}{a} + \frac{1}{2a^2(a^2+1)}$$

Differentiating repeatedly we can compute the integrals $\int_0^1 \frac{dx}{(x^2+a^2)^n}$

for $n=3, 4, \dots$. Their direct computation without using Leibniz' rule involves more lengthy calculations.

Up till now we considered integrals with constant limits of integration. But there also occur integrals with variable limits of integration which themselves are functions of the parameter the integrand depends on. For instance, this is the case in the inner integral involved into the repeated integral to which a double integral is reduced.

Let

$$F(\lambda) = \int_{a(\lambda)}^{b(\lambda)} f(x, \lambda) dx$$

We shall assume that the integrand $f(x, \lambda)$ and its partial derivative $f'_\lambda(x, \lambda)$ are continuous in a rectangle $a \leq x \leq b$, $\lambda_1 \leq \lambda \leq \lambda_2$ ($a, b, \lambda_1, \lambda_2 = \text{const}$), and that the functions $a(\lambda)$ and $b(\lambda)$ are continuous in the interval $[\lambda_1, \lambda_2]$, possess continuous derivatives and satisfy the conditions $a \leq a(\lambda) \leq b(\lambda) \leq b$. Then the function $F(\lambda)$ is continuous and differentiable, and its derivative is given by the formula

$$F'(\lambda) = \int_{a(\lambda)}^{b(\lambda)} f'_\lambda(x, \lambda) dx + f[b(\lambda), \lambda] b'(\lambda) - f[a(\lambda), \lambda] a'(\lambda)$$

(this is the *general form of Leibniz' rule*).

We shall not present the proof of this formula and confine ourselves to the remark that Leibniz' rule for integrals with constant

For instance, the integrals $\int_0^{\infty} \frac{\sin \lambda x}{1+x^2} dx$ and $\int_0^{\infty} \frac{\cos \lambda x}{1+x^2} dx$ satisfy the conditions of this definition for all the values of λ since the moduli of their integrands do not exceed the positive function $\frac{1}{1+x^2}$, the improper integral of the latter being convergent.

Regularly convergent improper integrals (*) possess the following properties which we state without proof:

1. If $f(x, \lambda)$ is a continuous function for $x \geq a$, $\lambda \in [\lambda_1, \lambda_2]$ and the integral $F(\lambda) = \int_a^{\infty} f(x, \lambda) dx$ is regularly convergent the function $F(\lambda)$ depends continuously on the parameter λ .

2. If, in addition to the conditions stated in 1, the function $f(x, \lambda)$ possesses the continuous partial derivative $f'_\lambda(x, \lambda)$ and the integral of this derivative is also regularly convergent the function $F(\lambda)$ is differentiable and its derivative is given by the formula

$$F'(\lambda) = \frac{d}{d\lambda} \int_a^{\infty} f(x, \lambda) dx = \int_a^{\infty} f'_\lambda(x, \lambda) dx$$

For an improper integral of the form $\int_a^b f(x, \lambda) dx$ involving a discontinuous integrand $f(x, \lambda)$ the notion of regular convergence is introduced in a similar fashion.

As an example demonstrating the properties stated above let us consider the important integral

$$I(\lambda) = \int_0^{\infty} e^{-x^2} \cos \lambda x dx \quad (**)$$

encountered in the theory of equations of mathematical physics. By the way, for $\lambda=0$ this integral turns into the integral

$$I(0) = \int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2} \text{ computed in Sec. 131. Integral (**) is regu-}$$

larly convergent since $|e^{-x^2} \cos \lambda x| \leq e^{-x^2}$. On differentiating integral (**) we receive

$$I'(\lambda) = - \int_0^{\infty} x e^{-x^2} \sin \lambda x dx$$

Such a differentiation is a legitimate operation in this case because the latter integral is also regularly convergent. Indeed, we have

$|xe^{-x^2} \sin \lambda x| \leq xe^{-x^2}$, and the integral $\int_0^\infty xe^{-x^2} dx$ is convergent:

$\int_0^\infty xe^{-x^2} dx = -\frac{e^{-x^2}}{2} \Big|_0^\infty = \frac{1}{2}$. To evaluate the integral expressing $I'(\lambda)$ we integrate by parts putting $xe^{-x^2} dx = dv$ and $\sin \lambda x = u$. Then $v = -\frac{1}{2}e^{-x^2}$ and $du = \lambda \cos \lambda x$, and we obtain

$$I'(\lambda) = -\int_0^\infty xe^{-x^2} \sin \lambda x dx = \frac{1}{2} e^{-x^2} \sin \lambda x \Big|_0^\infty - \frac{\lambda}{2} \int_0^\infty e^{-x^2} \cos \lambda x dx$$

The expression $\frac{1}{2} e^{-x^2} \sin \lambda x$ vanishes for $x=0$ and $x=\infty$, and the latter integral is again equal to $I(\lambda)$. We thus arrive at the relation $I'(\lambda) = -\frac{\lambda}{2} I(\lambda)$ which can be rewritten as

$$\frac{I'(\lambda)}{I(\lambda)} = [\ln I(\lambda)]' = -\frac{\lambda}{2}$$

On integrating we find $\ln I(\lambda) = -\frac{\lambda^2}{4} + C$. For $\lambda=0$ the integral $I(0)$ is equal to $\frac{\sqrt{\pi}}{2}$ whence $C = \ln \frac{\sqrt{\pi}}{2}$. Finally, $\ln I(\lambda) = -\frac{\lambda^2}{4} + \ln \frac{\sqrt{\pi}}{2}$ and

$$I(\lambda) = \frac{\sqrt{\pi}}{2} e^{-\frac{\lambda^2}{4}}$$

II. The gamma function. An important example of a non-elementary function specified by an improper integral dependent on a parameter is the so-called *gamma function*

$$\Gamma(\lambda) = \int_0^\infty x^{\lambda-1} e^{-x} dx$$

also termed *Euler's integral of the second kind*. The gamma function is frequently used in various problems of mathematical analysis and its applications.

First of all, let us show that the function $\Gamma(\lambda)$ is defined and continuous for $\lambda > 0$. To this end, it is sufficient to prove that the integral specifying the gamma function is regularly convergent for every closed interval $[\lambda_1, \lambda_2]$ where $0 < \lambda_1 < \lambda_2 < \infty$.

Note that if $\lambda \geq 1$ the integrand $x^{\lambda-1} e^{-x}$ is continuous for $x \geq 0$; if $\lambda < 1$ this function has an infinite discontinuity at the point

$x=0$. Therefore it is convenient to split the integral specifying the gamma function into two:

$$\Gamma(\lambda) = \int_0^1 x^{\lambda-1} e^{-x} dx + \int_1^{\infty} x^{\lambda-1} e^{-x} dx$$

Let us begin with studying the first integral. For $0 < x \leq 1$ we have the inequalities $x^{\lambda-1} e^{-x} \leq x^{\lambda-1} \leq x^{\lambda_1-1}$ because $\lambda \geq \lambda_1$. The

integral $\int_0^1 x^{\lambda_1-1} dx = \frac{x^{\lambda_1}}{\lambda_1} \Big|_0^1$ is convergent for $\lambda_1 > 0$; hence, the first integral in the above sum is regularly convergent for $\lambda \geq \lambda_1 > 0$.

In the second integral we have $x \geq 1$, and therefore $x^{\lambda-1} e^{-x} \leq x^{\lambda_1-1} e^{-x}$ since $\lambda \leq \lambda_2$. As $x \rightarrow \infty$ the function x^{λ_1-1} increases slower than the exponential function $e^{\alpha x}$ where α is any fixed positive number. Indeed, applying L'Hospital's rule we obtain $\lim_{x \rightarrow \infty} \frac{x^{\lambda_1-1}}{e^{\alpha x}} = 0$,

and hence the ratio $\frac{x^{\lambda_1-1}}{e^{\alpha x}}$ is bounded in the interval $1 \leq x < \infty$.

Let us choose $\alpha = \frac{1}{2}$, then $x^{\lambda_1-1} e^{-x} \leq M e^{\frac{x}{2}} e^{-x} = M e^{-\frac{x}{2}}$. Since the integral $\int_1^{\infty} M e^{-\frac{x}{2}} dx$ is convergent the second integral is regularly convergent for $\lambda \leq \lambda_2$.

Thus, each of the integrals in question is a continuous function and, consequently, the gamma function (which is their sum) is continuous in any closed interval $[\lambda_1, \lambda_2]$ where $0 < \lambda_1 < \lambda_2 < \infty$ (we can simply say that it converges for $\lambda > 0$). It can also be shown that the function $\Gamma(\lambda)$ possesses derivatives of any order.

Let us write down the integral expressing $\Gamma(\lambda + 1)$ and perform integration by parts by putting $x^\lambda = u$ and $e^{-x} dx = dv$:

$$\Gamma(\lambda + 1) = \int_0^{\infty} x^\lambda e^{-x} dx = -x^\lambda e^{-x} \Big|_0^{\infty} + \int_0^{\infty} \lambda x^{\lambda-1} e^{-x} dx$$

The expression $x^\lambda e^{-x}$ vanishes for $x=0$ and $x=\infty$; taking the factor λ outside the sign of integration in the integral on the right-hand side we arrive at the recurrence formula

$$\Gamma(\lambda + 1) = \lambda \Gamma(\lambda)$$

expressing the most important property of the gamma function.

Using the latter formula we can readily compute the value of the gamma function when λ is a positive integer. Noting that

$\Gamma(1) = \int_0^{\infty} e^{-x} dx = 1$ and applying the recurrence formula we obtain

$$\begin{aligned}\Gamma(n+1) &= n\Gamma(n) = n(n-1)\Gamma(n-1) = \dots \\ &\dots = n(n-1)(n-2) \dots 2 \cdot 1 \cdot \Gamma(1) = n!\end{aligned}$$

Thus, for any positive integer n we have

$$\Gamma(n+1) = n!$$

Let us also compute the value $\Gamma\left(\frac{1}{2}\right)$ frequently encountered in applications. Making the substitution $x = t^2$ in the integral expressing $\Gamma\left(\frac{1}{2}\right)$ we receive

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} x^{-\frac{1}{2}} e^{-x} dx = 2 \int_0^{\infty} e^{-t^2} dt = \sqrt{\pi}$$

$\left(\int_0^{\infty} e^{-t^2} dt \text{ is the Euler-Poisson integral evaluated in Sec. 131}\right)$.

Applying the recurrence formula we consecutively find

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}, \quad \Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi}, \quad \Gamma\left(\frac{7}{2}\right) = \frac{15}{8}\sqrt{\pi}, \dots$$

The values of the integral specifying the gamma function for other values of λ can be found by means of approximate methods; in special handbooks there are extensive tables of the values of the gamma function (e.g. see [14]).

There are many important integrals expressible in terms of the gamma function. An example of this type is *Euler's integral of the first kind*:

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$$

It can be shown (but we do not present the proof here) that the function $B(p, q)$ dependent on the two variables p and q is defined for $p > 0$, $q > 0$ and is expressed in terms of the gamma function by the formula

$$B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)} \quad (p > 0, q > 0)$$

Euler's integral of the first kind $B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx$ is also termed the *beta function*.

QUESTIONS

1. How is the volume of a curvilinear cylinder found?
2. What is the double integral of a given function over a given domain?
3. State the existence theorem for the double integral.
4. Enumerate the properties of the double integral.
5. Derive the rule for computing a double integral over a domain whose boundary has at most two common points with every straight line parallel to the one of coordinate axes. Write down the corresponding formulas.
6. State the rule for change of variables in the double integral.
7. What do we call a Jacobian? What is the element of area in curvilinear coordinates?
8. Derive the formula for the element of area in polar coordinates proceeding from geometrical considerations.
9. How do we compute a double integral in polar coordinates by means of repeated integration?
10. How is the mass of a nonhomogeneous plate with given density found?
11. Derive the formulas for computing the static moments and the moments of inertia of a plate.
12. How do we determine the mass of a physical body with given mass density?
13. What is the triple integral of a given function over a given spatial domain?
14. Enumerate the properties of the triple integral.
15. How are triple integrals computed in Cartesian coordinates?
16. How are triple integrals computed in cylindrical coordinates?
17. How are triple integrals computed in spherical coordinates?
18. Derive the formulas for computing the coordinates of the centre of gravity of a spatial body.
19. Derive the formulas for the moments of inertia of a body.
- 20*. State the definition of an integral dependent on a parameter. Give examples.
- 21*. State and prove the theorem of the continuity of an integral dependent on a parameter.
- 22*. State and prove the theorem on the differentiability of an integral with respect to the parameter it depends on (Leibniz' rule).
- 23*. What is a regularly convergent improper integral dependent on a parameter? Enumerate the properties of such integrals.
- 24*. What is the gamma function? Prove that it is continuous and enumerate its properties.

Chapter IX

LINE INTEGRALS AND SURFACE INTEGRALS. FIELD THEORY

§ 1. Line Integrals

138. The Work of a Field of Force. The Line Integral. In Sec. 86, II we considered the problem of computing the work of a variable force in a rectilinear motion of a body, the force being directed along the path of motion. Now we proceed to a more general problem of this kind.

Suppose that there is a *plane force field* defined in a domain D in the plane Oxy . This means that the material point placed in the domain D is subjected to the action of a force¹ F in the plane Oxy whose magnitude and direction are specified at every point of D (spatial fields will be treated in Sec. 146). The domain D may coincide with the entire plane Oxy or be a part of the plane.

The dependence of the force F on the location of its point of application is described by a relation

$$F = F(x, y)$$

where x and y are the coordinates of the point of application of the force. It is supposed that the field is *stationary*, that is, the force F does not depend on time t and is only specified by the coordinates of the point.

To specify the force $F(x, y)$ we can specify its projections on the coordinate axes which are also functions of the variables x

¹ In physics, to characterize a force field irrespective of the material point placed in it the notion of the *intensity* of the field is introduced which is understood as a vector representing the force of action of the field upon a "unit body": upon a body of unit mass if we deal with the field of gravitation or upon a particle carrying unit electric charge if we deal with an electrostatic field and the like. Then the force with which the field acts upon an arbitrary body is equal to the product of the intensity of the field by, respectively, the mass or the charge of the body, etc. It should also be noted that a field of force is a special case of a *vector field* (see Sec. 151).

and y . Denoting the projections as $P(x, y)$ and $Q(x, y)$ we write

$$\mathbf{F} = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$$

Let the material point move along a path L in the domain D under the action of the force field (see Fig. 188 where L is the line joining the points B and C). Let us compute the work performed by the force field $\mathbf{F}(x, y)$ in this motion of the point.

To this end, we break up the path L into n parts by means of points $B_1, B_2, B_3, \dots, B_n, B_{n+1}$ (the point B_1 coincides with B and B_{n+1} with C), the coordinates of the points of division B_k ,

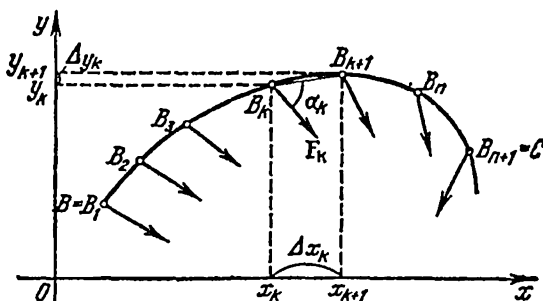


Fig. 188

$k = 1, 2, \dots, n+1$ being denoted x_k and y_k . Next we replace each curvilinear part $B_k B_{k+1}$ by the displacement vector $\overline{B_k B_{k+1}}$ and consider the force acting on the point as it passes the line segment $B_k B_{k+1}$ as, approximately, being constant and equal to the value of the force of the field at the point B_k :

$$\mathbf{F}_k = P(x_k, y_k)\mathbf{i} + Q(x_k, y_k)\mathbf{j}$$

Then the work corresponding to this part is

$$\Delta A_k = |\mathbf{F}_k| \cdot |\overline{B_k B_{k+1}}| \cos \alpha_k$$

where α_k is the angle between the vectors $\mathbf{F}_k$ and $\overline{B_k B_{k+1}}$. This expression is the scalar product of the vectors $\mathbf{F}_k$ and $\overline{B_k B_{k+1}}$, i.e.

$$\Delta A_k = \mathbf{F}_k \cdot \overline{B_k B_{k+1}}$$

The projections of the vector $\overline{B_k B_{k+1}}$ on the axes Ox and Oy being equal, respectively, to $x_{k+1} - x_k = \Delta x_k$ and $y_{k+1} - y_k = \Delta y_k$, we have

$$\overline{B_k B_{k+1}} = \Delta x_k \mathbf{i} + \Delta y_k \mathbf{j}$$

The well-known formula for the scalar product now implies the following expression for the work of the field on this part of the path of motion:

$$\Delta A_k = P(x_k, y_k) \Delta x_k + Q(x_k, y_k) \Delta y_k$$

Summing these expressions over all the parts the original path is divided into, we obtain an approximation to the sought-for work:

$$A_n = \sum_{k=1}^n P(x_k, y_k) \Delta x_k + Q(x_k, y_k) \Delta y_k \quad (*)$$

Finally, passing to the limit as $n \rightarrow \infty$ (on condition that the length of the longest part tends to zero) we obtain the true work of the field.

There are various problems leading to the passage to the limit in sums of type (*), not only the problem of work, and therefore

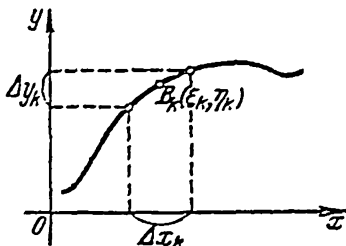


Fig. 189

we shall consider this question in the general form abstracting from the physical meaning of the functions $P(x, y)$ and $Q(x, y)$ as projections of a force. Let $P(x, y)$ and $Q(x, y)$ be arbitrary functions of two variables continuous in a domain D and let L be a smooth curve lying entirely within this domain. Dividing the curve L into n parts, choosing an arbitrary point $B_k(\xi_k, \eta_k)$ in each

segment of the curve (in particular, these points may coincide with the end points of the parts, as shown in Fig. 188) and denoting the projections of the k th part on the axes of coordinates by $\Delta x_k, \Delta y_k$ (Fig. 189) we form the sum

$$\sum_{k=1}^n P(\xi_k, \eta_k) \Delta x_k + Q(\xi_k, \eta_k) \Delta y_k \quad (**)$$

Sum (**) is called an *integral sum*, and its limit

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n P(x_k, y_k) \Delta x_k + Q(x_k, y_k) \Delta y_k = \int_L P(x, y) dx + Q(x, y) dy$$

is a *line integral over the path L* .¹

Definition. The limit of integral sum (**) as the length of the longest of the parts the line L is divided into tends to zero is called the *line integral over the path L* .

¹ An integral of this kind is sometimes termed a *line integral with respect to coordinates* or a *line integral of the second type*. Line integrals of another kind known as *line integral with respect to arc length* or *line integral of the first type* will be studied in Sec. 145.

The curve L is then called the *path* or the *contour of integration*, the point B is called the *initial point* and the point C the *terminal point of the path of integration*.

In particular, if $Q(x, y) \equiv 0$ the integral has the form

$$\int_L P(x, y) dx$$

and is called a *line integral with respect to the coordinate x* . Similarly, if $P(x, y) \equiv 0$ we speak about a *line integral with respect to the coordinate y* :

$$\int_L Q(x, y) dy$$

Coming back to the problem connected with the computation of work we can say that the work of a field of force performed as the material point moves in the contour L is expressed by the line integral

$$\int_L P(x, y) dx + Q(x, y) dy$$

where $P(x, y)$ and $Q(x, y)$ are the projections of the forces of the field on the coordinate axes.

139. Evaluating Line Integrals. Integral over a Closed Contour. We shall show that the evaluation of line integrals reduces to the computation of ordinary definite integrals. For instance, let us consider an integral of the form

$$\int_L P(x, y) dx$$

Suppose that the path of integration L is specified by parametric equations

$$x = x(t), \quad y = y(t)$$

where the functions $x(t)$, $y(t)$ and their derivatives are continuous. Let t_B be the value of the parameter t corresponding to the initial point B of the contour and t_C that corresponding to the terminal point C . Consider the integral sum

$$\sum_{k=1}^n P(\xi_k, \eta_k) \Delta x_k$$

whose limit is the given line integral. Let us express it in terms of the variable t . Since

$$\Delta x_k = x_{k+1} - x_k = x(t_{k+1}) - x(t_k)$$

the application of Lagrange's formula (see Sec. 57) gives us

$$\Delta x_k = x'(\theta_k) \Delta t_k$$

where θ_k belongs to the interval $[t_k, t_{k+1}]$ and $\Delta t_k = t_{k+1} - t_k$. As the intermediate point (ξ_k, η_k) we shall take the one corresponding to the value θ_k of the parameter. Then

$$\xi_k = x(\theta_k), \quad \eta_k = y(\theta_k)$$

The transformed expression

$$\sum_{k=1}^n P[x(\theta_k), y(\theta_k)] x'(\theta_k) \Delta t_k$$

is an ordinary integral sum associated with the function of one variable $P[x(t), y(t)] x'(t)$, and its limit is nothing but the definite integral

$$\int_{t_B}^{t_C} P[x(t), y(t)] x'(t) dt$$

Thus,

$$\int_L P(x, y) dx = \int_{t_B}^{t_C} P[x(t), y(t)] x'(t) dt$$

We similarly conclude that

$$\int_L Q(x, y) dy = \int_{t_B}^{t_C} Q[x(t), y(t)] y'(t) dt$$

What has been said implies the following *computation rule for the line integral*:

To transform the line integral

$$\int_L P(x, y) dx + Q(x, y) dy$$

taken over the curve $x=x(t)$, $y=y(t)$ to an ordinary definite integral one should replace x , y , dx and dy in the element of integration by their expressions in terms of t and dt and evaluate the resulting definite integral over the interval of variation of the parameter t :

$$\begin{aligned} & \int_L P(x, y) dx + Q(x, y) dy = \\ & = \int_{t_B}^{t_C} \{P[x(t), y(t)] x'(t) + Q[x(t), y(t)] y'(t)\} dt \end{aligned}$$

where t_B and t_C are the values of the parameter t corresponding to the points B and C of the curve L .

It follows from this rule that the line integral always exists if the functions $P(x, y)$ and $Q(x, y)$ are continuous and the functions

$x(t)$ and $y(t)$ are continuous together with their derivatives. The latter condition means (if we assume additionally that the derivatives $x'(t)$ and $y'(t)$ do not vanish simultaneously) that L is a *smooth* curve, that is, it has a tangent whose position varies continuously. By the way, the line L may consist of several arcs for each of which this condition holds. Then the whole curve L is termed *piecewise smooth*, and, for continuous functions P and Q , the line integral over such a curve also exists. When an integral over a piecewise smooth contour L is computed the latter should be divided into smooth parts, after which the integrals over these parts are computed and added together.

If the line L is specified by an equation of the form $y = y(x)$ we can use the general formula by putting $t = x$ to obtain

$$\int_L P(x, y) dx + Q(x, y) dy = \int_{x_B}^{x_C} \{P[x, y(x)] + Q[x, y(x)] y'(x)\} dx$$

where x_B and x_C are the abscissas of the points B and C .

Let the reader write down the reduction formula for a line integral to the definite integral when the path of integration is determined by an equation $x = x(y)$.

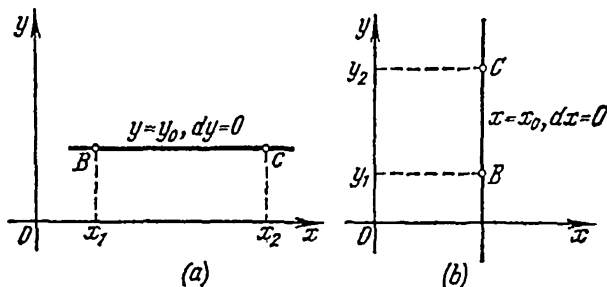


Fig. 190

If the whole contour of integration consists of a number of arcs specified by equations of different types the contour should be divided into parts and the desired integral should be computed as the sum of the integrals taken over the parts.

We also mention the following two important special cases:

If the contour of integration L is a line segment parallel to the axis of abscissas (see Fig. 190a) the line integral is originally written as a definite integral. Indeed, in this case we have $y = y_0$ and $dy = 0$, and hence

$$\int_L P(x, y) dx + Q(x, y) dy = \int_{x_1}^{x_2} P(x, y_0) dx$$

If L is a segment of a straight line parallel to the axis of ordinates (Fig. 190b), we have $x = x_0$, $dx = 0$, and hence

$$\int_L P(x, y) dx + Q(x, y) dy = \int_{y_1}^{y_2} Q(x_0, y) dy$$

Examples. (1) Let us evaluate the integral $I = \int_L xy dx + (x + y) dy$ if the line L (see Fig. 191) is:

(a) the line segment joining the point $O(0, 0)$ to the point $A(1, 1)$;

(b) the arc of the parabola $y = x^2$ connecting the same points;

(c) the polygonal line OBA .

In Case (a) the equation of the path of integration is $y = x$. Consequently, $dy = dx$, and we obtain

$$I = \int_0^1 (x^2 + 2x) dx = \left(\frac{x^3}{3} + x^2 \right) \Big|_0^1 = \frac{4}{3}$$

In Case (b) we have $y = x^2$ and $dy = 2x dx$, and thus

$$I = \int_0^1 [x^3 + (x + x^2) 2x] dx = \frac{17}{12}$$

In Case (c) the contour of integration should be split into the two parts OB and BA . For the part OB we have $y = 0$ and $dy = 0$ and for BA we have $x = 1$ and $dx = 0$. Therefore

$$I = \int_0^1 (1 + y) dy = \left(\frac{1 + y^2}{2} \right) \Big|_0^1 = \frac{3}{2}$$

We see that in all the three cases we have obtained *different* results although the initial and the terminal points of the contour of integration remained invariable. Hence, generally speaking, the line integral depends not only on the initial and terminal points of the contour of integration but also on the shape of the curve connecting these points. This question will be discussed in more detail in Sec. 141.

(2) Let us compute the integral

$$I = \int_L xy dx$$

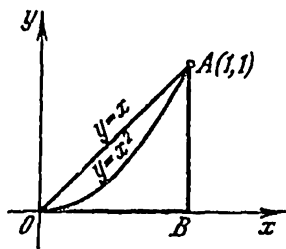


Fig. 191

of integration expressing x as a one-valued function of y , we substitute $x=y^2$ and $dx=2ydy$ into the integral and receive

$$I = \int_{-1}^1 y^2 y \cdot 2y dy = \frac{2y^5}{5} \Big|_{-1}^1 = \frac{4}{5}$$

If the same contour L is specified so that x is expressed as a function of y it should be split into the two parts AO and OB whose equations are $y = -\sqrt{x}$ and $y = +\sqrt{x}$. Then we obtain

$$\begin{aligned} I &= \int_{AO} xy dx + \int_{OB} xy dx = \\ &= -\int_1^0 x \sqrt{x} dx + \int_0^1 x \sqrt{x} dx = \\ &= 2 \int_0^1 x^{\frac{3}{2}} dx = \frac{4}{5} \end{aligned}$$

(3) Let us find the integral

$$I = \int_L y dx - x dy$$

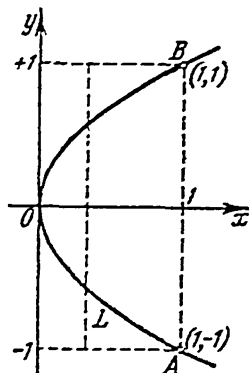


Fig. 192

where L is the arc of the cycloid $x = 2(t - \sin t)$, $y = 2(1 - \cos t)$ joining its points $O(0, 0)$ and $A(4\pi, 0)$. Using the general rule and taking into account that the value of the parameter corresponding to the point O is $t=0$ and that corresponding to A is $t=2\pi$ we obtain

$$\begin{aligned} I &= \int_0^{2\pi} [4(1 - \cos t)^2 - 4(t - \sin t) \sin t] dt = \\ &= 4 \int_0^{2\pi} (2 - 2 \cos t - t \sin t) dt = 4 \left[2t \Big|_0^{2\pi} - 2 \sin t \Big|_0^{2\pi} - \int_0^{2\pi} t \sin t dt \right] \end{aligned}$$

Integration by parts in the last integral leads to the final result:

$$I = 16\pi - 4(-t \cos t + \sin t) \Big|_0^{2\pi} = 24\pi$$

Up till now we dealt with nonclosed contours of integration whose initial and terminal points were different. The specification of these points determined the *direction of integration*. It is clear that *if the direction of integration is reversed the line integral changes its sign to the opposite*. This conclusion can be drawn, for instance, from the general rule for reduction of the line

integral to the definite one. Writing, for brevity, P and Q instead of $P(x, y)$ and $Q(x, y)$ we express this property as

$$\int_{-L} P dx + Q dy = - \int_{+L} P dx + Q dy$$

where the symbols $-L$ and $+L$ designate (conditionally) the line L taken with the two opposite directions.

There are also many cases when a line integral is taken over a *closed contour*. We shall limit ourselves to the case when the contour is a piecewise smooth closed *simple curve* (i.e. not intersecting itself) of the type shown in Fig. 193. Line integrals of this kind are designated by the symbol $\oint_L$. It is obvious that for

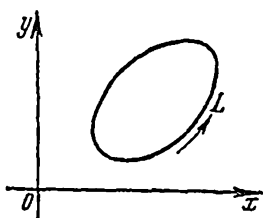


Fig. 193

an integral over a closed contour the choice of the initial point (which simultaneously is the terminal point) is inessential, and the value of the integral depends solely on the *orientation* of the contour, that is, on the choice of the direction of integration.

For a (simple) closed contour in the plane the direction in which it is traversed is termed *positive* if the domain bounded by the contour always remains on the left. The opposite direction is called *negative* (we also say that the contour is *oriented* positively or negatively). In the former case the contour is traced counterclockwise and in the latter clockwise. In this terminology it is supposed that the orientations of the contour and of the coordinate system are coherent: if the positively oriented contour is traced counterclockwise the shortest rotation from Ox to Oy should also be in the counterclockwise direction as shown in Fig. 193 (a right-handed coordinate system).

Let us agree that in what follows an integral $\int_L P dx + Q dy$ over a closed contour L is understood as being taken in the positive direction; the integration in the opposite direction will be symbolized as

$$\int_{-L} P dx + Q dy$$

Our discussion shows that these integrals only differ in sign. The following proposition indicates a specific property of line integrals over closed contours:

Theorem. If a domain D bounded by a closed contour L is split into two parts D_1 and D_2 the line integral over the whole contour L is equal to the sum of the integrals taken in the same

direction over the contours L_1 and L_2 bounding the domains D_1 and D_2 .

Proof. Let the domain D be bounded by the closed curve L shown as the line $AECBA$ in Fig. 194 and the domains D_1 and D_2 by the lines L_1 and L_2 shown, respectively, as $AECA$ and $ACBA$. We apparently have

$$\int_{L_1} P dx + Q dy = \int_{AEC} P dx + Q dy + \int_{CA} P dx + Q dy$$

and

$$\int_{L_2} P dx + Q dy = \int_{AC} P dx + Q dy + \int_{CBA} P dx + Q dy$$

The integrals over CA and AC are taken along one and the same line but in the opposite directions, and therefore their sum is equal to zero. Finally, on adding the last two equalities termwise we obtain

$$\begin{aligned} \int_{L_1} P dx + Q dy + \int_{L_2} P dx + Q dy &= \\ &= \int_{AEC} P dx + Q dy + \int_{CBA} P dx + Q dy = \\ &= \int_L P dx + Q dy \end{aligned}$$

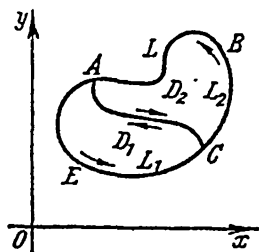


Fig. 194

which is what we set out to prove.

This theorem obviously remains true for any number of subdomains the domain D is split into.

140. Green's¹ Formula. In this section we shall prove *Green's theorem* which provides a formula connecting a line integral over a closed contour with a double integral over the domain bounded by the contour. This formula will play an important role.

Theorem. If the functions $P(x, y)$ and $Q(x, y)$ are continuous together with their partial derivatives in a closed bounded domain D then

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_L P dx + Q dy \quad (*)$$

where L is the boundary of the domain D and the integration along L goes in the positive direction.²

¹ Green, G. (1793-1841), an English physicist and mathematician.

² More precisely, we shall suppose that the functions P and Q and their partial derivatives are defined and continuous in a wider open domain G containing the closed domain D . In what follows we shall essentially use this condition.

Relation (*) is known as *Green's formula*.

Proof. Let us begin with a domain D in the plane Oxy bounded by a curve L having at most two common points with every straight line parallel to Ox or to Oy (Fig. 195). We shall transform the double integral $I =$

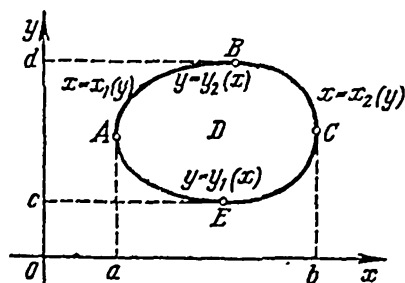


Fig. 195

$= \iint_D \frac{\partial P(x, y)}{\partial y} dx dy$ by integrating first with respect to y and then with respect to x , which yields

$$I = \int_a^b dx \int_{y_1(x)}^{y_2(x)} \frac{\partial P(x, y)}{\partial y} dy$$

where $y = y_2(x)$ is the equation of the arc ABC and $y = y_1(x)$ that of the arc AEC . On performing the inner integration we obtain

$$\begin{aligned} I &= \int_a^b P(x, y) \Big|_{y_1(x)}^{y_2(x)} dx = \int_a^b \{P[x, y_2(x)] - P[x, y_1(x)]\} dx = \\ &= \int_a^b P[x, y_2(x)] dx + \int_b^a P[x, y_1(x)] dx \end{aligned}$$

(we interchanged the limits of integration in the last integral which resulted in the change of sign).

The first integral is nothing but the line integral $\int_{ABC} P(x, y) dx$, and the second is the line integral $\int_{CEA} P(x, y) dx$. The sum of these integrals is the line integral of P with respect to x taken round the whole contour L in the *negative* direction. Consequently,

$$\iint_D \frac{\partial P}{\partial y} dx dy = - \int_L P dx \quad (**)$$

Similarly,

$$\iint_D \frac{\partial Q}{\partial x} dx dy = \int_c^d dy \int_{x_1(y)}^{x_2(y)} \frac{\partial Q}{\partial x} dx = \int_c^d \{Q[x_2(y), y] - Q[x_1(y), y]\} dy$$

where $x = x_1(y)$ and $x = x_2(y)$ are, respectively, the equations of the curves EAB and ECB written in the form solved in x . Arguing as in the former case we conclude that

$$\iint_D \frac{\partial Q}{\partial x} dx dy = \int_{ECB} Q(x, y) dy + \int_{BAE} Q(x, y) dy$$

The sum of the integrals on the right-hand side of the last equality can be replaced by one integral over the whole contour L this time taken in the *positive* direction:

$$\iint_D \frac{\partial Q}{\partial x} dx dy = \int_L Q dy \quad (***)$$

Finally, on subtracting termwise equality (**) from equality (***) we receive

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_L P dx + Q dy$$

which is what we had to prove.

Formula (**) also remains valid in the case of a domain depicted in Fig. 196. In this case the computation of the double integral $\iint_D \frac{\partial P}{\partial y} dx dy$ leads (in the notation

of Fig. 196) to the sum of integrals $\int_{A_2 C_2} P dx + \int_{C_1 A_1} P dx$. The integrals $\int_{C_2 C_1} P dx$ and $\int_{A_1 A_2} P dx$ are equal to zero since $C_2 C_1$

and $A_1 A_2$ are vertical line segments for which $x = \text{const}$, $dx = 0$. Adding these integrals to the above sum we again

arrive at the integral $-\int_L P dx$ where L is the contour $A_1 C_1 C_2 A_2$.

Similarly, formula (***) is also applicable to domains whose boundaries involve line segments parallel to Ox .

If a domain D bounded by a contour L is of a more complicated shape than those discussed above, it can usually be divided into a finite number of subdomains to which formulas (**) and (***) apply. Hence, we can write down Green's formula for each of the subdomains. Then, if we add together the resulting equalities the sum of the double integrals is equal to the double integral over the entire domain D while, according to the theorem proved in Sec. 139, the sum of the line integrals over the contour bounding the subdomains is equal to the line integral over the contour L . Therefore Green's formula is valid for a domain of this kind as well. For the general case, when it is only known that the contour L bounding the given domain D is a piecewise smooth simple closed curve, the proof involves more sophisticated considerations, and we shall not present them here.

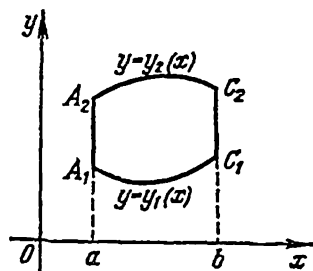


Fig. 196

141. Conditions for a Line Integral Being Path-independent. As was mentioned in the solution of Example 1, Sec. 139, for a given integrand and given initial and terminal points, the line integral

$$\int_L P(x, y) dx + Q(x, y) dy \quad (*)$$

may depend on the shape of the curve L joining these points. In this connection it is natural to pose the question what are the conditions guaranteeing the independence of integral $(*)$ of

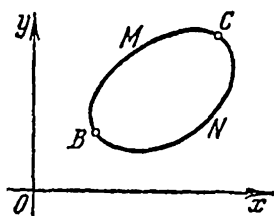


Fig. 197

the choice of the path of integration and its dependence only on the initial and terminal points of the path? When the latter property takes place we say that the line integral is *path-independent*.

From the point of view of mechanics, the path-independence of integral $(*)$ in the case when $P(x, y)$ and $Q(x, y)$ are the projections of a force field $F = P(x, y) i + Q(x, y) j$ means that the work of this field does not depend on the shape of the

trajectory of motion and depends solely on its initial and terminal points. As is known from physics, the work performed in the motion in the field of gravitation is independent of the form of the path; the solution of the general question posed here makes it possible to investigate this problem for other, more complicated, force fields.

We shall begin with proving a simple auxiliary proposition which allows us to reduce this question to another problem whose solution can be obtained on the basis of the results of Sec. 140.

For the sake of brevity, we shall denote the integral $\int_L P(x, y) dx + Q(x, y) dy$ by the symbol I_L .

Lemma. For line integral $(*)$ to be independent of the path of integration it is necessary and sufficient that this integral taken over any closed contour be equal to zero.

Proof. Suppose that it is known that integral is path-independent; we shall prove that this integral, when taken over any closed contour, is equal to zero. Let us choose an arbitrary closed contour L . Mark two points B and C on it (see Fig. 197). Since, by the hypothesis, the integral along the arc BMC is equal to that along the arc BNC , i.e.

$$I_{BMC} = I_{BNC}$$

the integral over the whole contour L is

$$I_L = I_{BNC} + I_{CMB} = I_{BNC} - I_{BMC} = 0$$

Conversely, let it be known that integral (*) turns into zero for any closed contour; we must show that in this case it is path-independent. To this end, we choose two arbitrary curves BMC and BNC joining two given points B and C (see again Fig. 197). According to the hypothesis, the integral over the closed contour $BNCMB$ is equal to zero:

$$I_{BNCMB} = 0, \text{ i.e. } I_{BNC} + I_{CMB} = 0$$

It now follows that

$$I_{BMC} = -I_{CNB} = I_{BNC}$$

The lemma has thus been completely proved.

Now we proceed to the basic theorems of our investigation (before studying these theorems the reader should recall the definition of a simply connected domain stated in Sec. 107, III and the subject matter of Sec. 115).

Theorem I. Let the functions $P(x, y)$ and $Q(x, y)$ be continuous together with their partial derivatives in a simply connected domain D . Then, for the line integral

$$\int_L P(x, y) dx + Q(x, y) dy \quad (*)$$

to be independent of the path of integration lying within the domain D it is necessary and sufficient that for all the points of the domain D the condition

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (**)$$

should hold.

If the functions P and Q possess these properties for any x and y , the domain D coincides with the entire plane Oxy .

Proof. According to the above lemma, the condition that integral (*) is path-independent is equivalent to the condition that it turns into zero for any closed contour lying entirely within the domain D .

Let us prove that condition (**) is *sufficient* for integral (*) over any closed contour to be equal to zero. Take an arbitrary closed contour L^* bounding a domain D^* and lying entirely inside the domain D (see Fig. 198a). Applying Green's formula to this contour we obtain

$$\int_{L^*} P dx + Q dy = \iint_{D^*} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$$

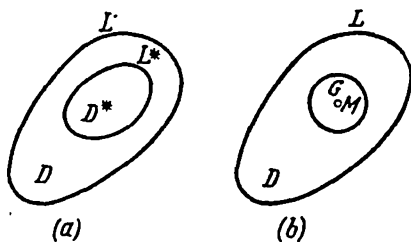


Fig. 198

141. Conditions for a Line Integral Being Path-independent. As was mentioned in the solution of Example 1, Sec. 139, for a given integrand and given initial and terminal points, the line integral

$$\int_L P(x, y) dx + Q(x, y) dy \quad (*)$$

may depend on the shape of the curve L joining these points. In this connection it is natural to pose the question what are the conditions guaranteeing the independence of integral $(*)$ of the choice of the path of integration and its dependence only on the initial and terminal points of the path? When the latter property takes place we say that the line integral is *path-independent*.

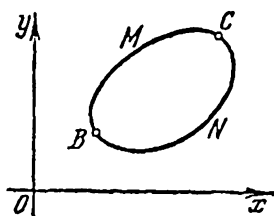


Fig. 197

From the point of view of mechanics, the path-independence of integral $(*)$ in the case when $P(x, y)$ and $Q(x, y)$ are the projections of a force field $F = P(x, y) i + Q(x, y) j$ means that the work of this field does not depend on the shape of the

trajectory of motion and depends solely on its initial and terminal points. As is known from physics, the work performed in the motion in the field of gravitation is independent of the form of the path; the solution of the general question posed here makes it possible to investigate this problem for other, more complicated, force fields.

We shall begin with proving a simple auxiliary proposition which allows us to reduce this question to another problem whose solution can be obtained on the basis of the results of Sec. 140.

For the sake of brevity, we shall denote the integral $\int_L P(x, y) dx + Q(x, y) dy$ by the symbol I_L .

Lemma. For line integral $(*)$ to be independent of the path of integration it is necessary and sufficient that this integral taken over any closed contour be equal to zero.

Proof. Suppose that it is known that integral is path-independent; we shall prove that this integral, when taken over any closed contour, is equal to zero. Let us choose an arbitrary closed contour L . Mark two points B and C on it (see Fig. 197). Since, by the hypothesis, the integral along the arc BMC is equal to that along the arc BNC , i.e.

$$I_{BMC} = I_{BNC}$$

the integral over the whole contour L is

$$I_L = I_{BNC} + I_{CMB} = I_{BNC} - I_{BMC} = 0$$

Conversely, let it be known that integral (*) turns into zero for any closed contour; we must show that in this case it is path-independent. To this end, we choose two arbitrary curves BMC and BNC joining two given points B and C (see again Fig. 197). According to the hypothesis, the integral over the closed contour $BNCMB$ is equal to zero:

$$I_{BNCMB} = 0, \text{ i.e. } I_{BNC} + I_{CMB} = 0$$

It now follows that

$$I_{BMC} = -I_{CNB} = I_{BNC}$$

The lemma has thus been completely proved.

Now we proceed to the basic theorems of our investigation (before studying these theorems the reader should recall the definition of a simply connected domain stated in Sec. 107, III and the subject matter of Sec. 115).

Theorem I. Let the functions $P(x, y)$ and $Q(x, y)$ be continuous together with their partial derivatives in a simply connected domain D . Then, for the line integral

$$\int_L P(x, y) dx + Q(x, y) dy \quad (*)$$

to be independent of the path of integration lying within the domain D it is necessary and sufficient that for all the points of the domain D the condition

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (**)$$

should hold.

If the functions P and Q possess these properties for any x and y , the domain D coincides with the entire plane Oxy .

Proof. According to the above lemma, the condition that integral (*) is path-independent is equivalent to the condition that it turns into zero for any closed contour lying entirely within the domain D .

Let us prove that condition (**) is *sufficient* for integral (*) over any closed contour to be equal to zero. Take an arbitrary closed contour L^* bounding a domain D^* and lying entirely inside the domain D (see Fig. 198a). Applying Green's formula to this contour we obtain

$$\int_{L^*} P dx + Q dy = \iint_{D^*} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy$$

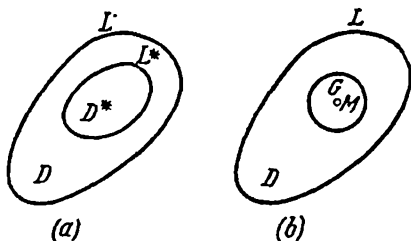


Fig. 198

Since, by the hypothesis, $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$, the double integral is equal to zero. Consequently, the line integral over the contour L^* is also equal to zero, which is what we set out to prove.

Now we proceed to prove the *necessity* of condition (**). Here we suppose that integral (*) over *any* closed contour is equal to zero. Then the application of Green's formula shows that the double integral of $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ over *any* domain G contained inside D is equal to zero:

$$\iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = 0 \quad (***)$$

Suppose that the equality $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y}$ does not hold. Let M be a point at which this equality is violated. Consider the difference $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ and suppose that it is positive at the point M :¹

$$\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)_M > 0$$

By the hypothesis that the partial derivatives are continuous, it follows that the difference $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ retains sign in a sufficiently small neighbourhood of the point M . Let us take an open domain G belonging to this neighbourhood and containing the point M (see Fig. 198b). Then we have

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} > 0$$

for all the points of G . Therefore the well-known property of double integrals (Sec. 129) implies that

$$\iint_G \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy > 0$$

which contradicts equality (***). This indicates that the assumption that condition (**) is violated is not true, that is, we have $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ for the entire domain D . The proof of the theorem has been completed.

The reader has undoubtedly recalled that the condition $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ involved in this theorem was dealt with in Sec. 115; as was said, in the case of a simply connected domain this condition is neces-

¹ It is seen from the course of the proof that the case $\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)_M < 0$ can be treated similarly.

sary for the expression $P(x, y)dx + Q(x, y)dy$ to be the total differential of a function $u(x, y)$. As to the sufficiency of this condition and the method for finding the function $u(x, y)$, they were only established for domains of a special class: rectangles with sides parallel to the coordinate axes and their extensions to unbounded domains of the type of a strip, a quadrant and a half-plane. Now we are going to elaborate the proof of the sufficiency theorem stated in Sec. 115 for the general case.

Theorem II. If the functions $P(x, y)$ and $Q(x, y)$ and their partial derivatives are continuous in a simply connected domain D and the equality $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ holds for all its points the expression $P(x, y)dx + Q(x, y)dy$ is a total differential.

Proof. Consider the line integral $\int_L P(x, y)dx + Q(x, y)dy$ along a curve L lying in the domain D and connecting a fixed point $M_0(x_0, y_0)$ to the variable point $M(x, y)$. According to Theorem I, this integral is path-independent and, for the fixed initial point M_0 , it depends solely on the terminal point $M(x, y)$, that is, on its coordinates x and y . Thus, this line integral is a function of the variables x and y , which we write as

$$I(x, y) = \int_{(x_0, y_0)}^{(x, y)} P(x, y)dx + Q(x, y)dy$$

The shape of the line joining the points M_0 and M is inessential here, and therefore it is not indicated in the notation.

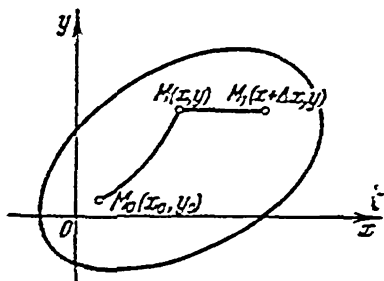


Fig. 199

Let us prove that the total differential of the function $I(x, y)$ is equal to the element of integration $Pdx + Qdy$ (this is analogous to the fact that the derivative of a definite integral with variable upper limit of integration with respect to that limit is equal to the integrand). Let us compute the derivative $I'_x(x, y)$. The partial increment $\Delta_x I(x, y)$ is given by the formula

$$\Delta_x I = I(x + \Delta x, y) - I(x, y) = \int_{(x_0, y_0)}^{(x + \Delta x, y)} Pdx + Qdy - \int_{(x_0, y_0)}^{(x, y)} Pdx + Qdy$$

If the second integral on the right is taken along an arbitrary line M_0M (see Fig. 199; the shape of this line is inessential) then the first integral is representable as the sum of the integral taken over the same line M_0M and the integral along the line

segment MM_1 lying on a straight line $y = \text{const.}$ This representation is allowable since both integrals do not depend on the path of integration. The integrals with opposite signs over the curve M_1M cancel out, and the latter equality takes the form

$$\Delta_x I = \int_{(x, y)}^{(x+\Delta x, y)} P(x, y) dx + Q(x, y) dy$$

In this integral the integration is carried out along a line segment parallel to the axis of abscissas and, as was mentioned on p. 487, the line integral reduces to an ordinary definite integral:

$$\Delta_x I = \int_x^{x+\Delta x} P(x, y) dx$$

where y is considered constant in the integration process. Applying to this integral the mean-value theorem and dividing both sides by Δx we get

$$\frac{\Delta_x I}{\Delta x} = \frac{P(\xi, y) \Delta x}{\Delta x} = P(\xi, y) \text{ where } \xi \in (x, x + \Delta x)$$

If $\Delta x \rightarrow 0$ the point (ξ, y) tends to the point (x, y) , and, since the function $P(x, y)$ is continuous, we have $P(\xi, y) \rightarrow P(x, y)$. Therefore the left-hand side of the equality also has a limit, as $\Delta x \rightarrow 0$, which is nothing but the partial derivative $I'_x(x, y)$:

$$\lim_{\Delta x \rightarrow 0} \frac{\Delta_x I}{\Delta x} = \frac{\partial I}{\partial x} = P(x, y)$$

The relation $I'_y(x, y) = Q(x, y)$ is proved completely analogously.

Thus, we have found the partial derivatives of the function $I(x, y)$; they are continuous, and hence, by the theorem in Sec. 111, the function $I(x, y)$ is differentiable, and its total differential is

$$dI = P(x, y) dx + Q(x, y) dy$$

which is what we intended to prove.

Now, summing up the results of our study, we can say that if the domain D is simply connected and the functions $P(x, y)$ and $Q(x, y)$ are continuous in this domain together with their partial derivatives the following four conditions are equivalent (that is, each of them implies the other three):

1. The line integral $\int_L P dx + Q dy$ taken round any closed contour lying entirely inside the domain D is equal to zero.

2. The line integral $\int_L P dx + Q dy$ does not depend on the path of integration joining two fixed points.

3. The expression $P(x, y) dx + Q(x, y) dy$ is a total differential.

4. At every point of the domain D the equality $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ is fulfilled.

The fourth condition is of particular importance since it is usually applied in practice for checking the other three conditions.

It was stressed that the domain D should be simply connected. It is natural to pose the question as to what situation occurs in the case of a multiply connected domain. It turns out that for a multiply connected domain the first three conditions remain equivalent to each other, and each of them implies the equality $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, but the converse does not necessarily hold since there are cases when the latter equality is fulfilled for all the points of a multiply connected domain while the other three conditions are violated. What has been said is confirmed by the following example.

Example. Consider the expression

$$-\frac{y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy = \frac{-y dx + x dy}{x^2+y^2}$$

Here we have

$$P = -\frac{y}{x^2+y^2}, \quad \frac{\partial P}{\partial y} = \frac{-x^2-y^2+2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

$$Q = \frac{x}{x^2+y^2}, \quad \frac{\partial Q}{\partial x} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

The functions $P(x, y)$ and $Q(x, y)$ and their partial derivatives are everywhere continuous except at the origin, and the equality $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ is fulfilled. Therefore in any simply connected domain not containing the origin the other three conditions are also fulfilled. In particular, if L is a closed contour not encircling the origin (and, of course, not passing through it) we have

$$I_L = \int_L \frac{-y dx + x dy}{x^2+y^2} = 0$$

The situation changes sharply if we take the annulus shown in Fig. 200 which is a doubly connected domain not containing the point $O(0, 0)$ (the radius r_1 can be arbitrarily small and the radius r_2 arbitrarily large). It is clear that the fourth condition holds throughout the annulus. Let a circle of radius R where

$r_1 < R < r_2$ be specified by the parametric equations $x = R \cos t$, $y = R \sin t$. The integral over this circle is equal to

$$\int_0^{2\pi} \frac{-R \sin t (-R \sin t) + R \cos t \cdot R \cos t}{R^2 \cos^2 t + R^2 \sin^2 t} dt = \int_0^{2\pi} dt = 2\pi$$

We thus see that there are closed contours within this annulus for which the integrals taken round them are nonzero. Fig. 200 also clearly indicates that $I_{L_1} + I_{-L_1} = 2\pi$, and, since $I_{-L_1} = -I_{L_1}$, we have

$$I_{L_2} = I_{L_1} + 2\pi$$

This means that the integrals over the contours L_1 and L_2 joining the points M_1 and M_2 have different values.

A more complicated situation arises in connection with the third assertion. It turns out that there is no single-valued function whose total differential is equal to the given expression. This can readily be shown if we pass to polar coordinates. Then we receive

$$x = r \cos \varphi, \quad y = r \sin \varphi \quad \text{and} \quad dx = \cos \varphi dr - r \sin \varphi d\varphi, \\ dy = \sin \varphi dr + r \cos \varphi d\varphi$$

and the given expression is transformed as

$$\frac{-y dx + x dy}{x^2 + y^2} = d\varphi$$

This shows that it is equal to the total differential of the polar angle φ as function of the Cartesian coordinates x and y , and the question is for what domains it is possible to separate one-valued branches of this function. As is well known, this function is not one-valued since for any given x and y it has infinitely many values differing from each other by integer multiples of 2π . If we take a simply connected domain not containing the origin it is possible to separate single-valued branches of the function φ . For instance, such a branch can be chosen by setting an interval of variation of φ of length 2π of the type $0 \leq \varphi < 2\pi$ or $-\pi < \varphi \leq \pi$, etc. But in the annulus we deal with it is impossible to construct a single-valued branch of the function φ . For, when the variable point M traverses counterclockwise a circle lying within the annulus and having its centre at the origin, the magnitude φ varies continuously and gains an increment of 2π as the point M returns to the initial position (this increment is

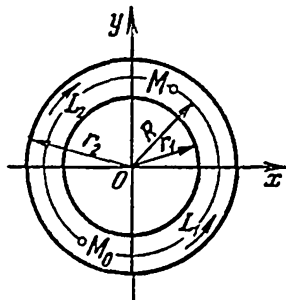


Fig. 200

exactly equal to the value of the integral of the given expression taken over that circle).

The questions discussed here are investigated more thoroughly in the theory of functions of a complex argument.

142. Integrating Total Differentials. Primitive. We shall consider the line integral

$$\int_L P(x, y) dx + Q(x, y) dy \quad (*)$$

under the assumption that all the conditions enumerated in Sec. 141 are fulfilled: the contour L belongs to a simply connected domain D in which the functions $P(x, y)$ and $Q(x, y)$ and their partial derivatives are continuous and the condition $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ holds. Then the value of integral (*) only depends on the position of the initial and terminal points of the path of integration; as was agreed, in this case we write the integral in the form

$$\int_{(x_0, y_0)}^{(x_1, y_1)} P(x, y) dx + Q(x, y) dy$$

where (x_0, y_0) are the coordinates of the initial point of the contour of integration and (x_1, y_1) those of the terminal point. Now, what is the most convenient way of computing such an integral? It may seem that the simplest way is to integrate along the line segment connecting these points but this is not so. We shall make use of the remark on pp. 487, 488 and take as the path of integration the broken line $M_0 M_2 M_1$ (or $M_0 M_3 M_1$) shown in Fig. 201 whose segments are parallel to the coordinate axes. Since we have $y = y_0$ and $dy = 0$ for the segment $M_0 M_2$ and $x = x_1$ and $dx = 0$ for the segment $M_2 M_1$ the integral is represented in the form

$$\int_{(x_0, y_0)}^{(x_1, y_1)} P(x, y) dx + Q(x, y) dy = \int_{x_0}^{x_1} P(x, y_0) dx + \int_{y_0}^{y_1} Q(x_1, y) dy$$

Let the reader write down a similar integration formula for the broken line $M_0 M_3 M_1$.

If it turns out that the broken lines $M_0 M_2 M_1$ and $M_0 M_3 M_1$ fall outside the domain D where the conditions of the basic theorem are fulfilled (see Fig. 202) we can choose another path of integration of this kind, for instance, such as the broken line $M_0 M_4 M_3 M_1$.

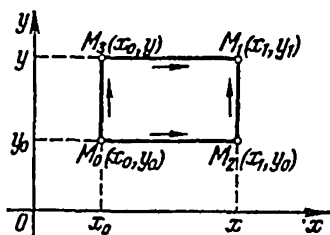


Fig. 201

Example. Let us compute the integral

$$I = \int_{(0,0)}^{(2,1)} 2xy \, dx + x^2 \, dy$$

In this example the element of integration turns into zero on any line segment lying on the axis of abscissas, and hence we have

$$I = \int_0^1 2^2 dy = 4$$

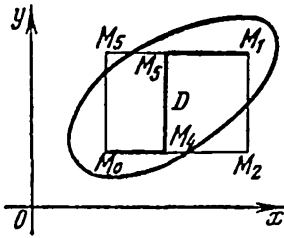


Fig. 202

Now let us consider an integral with fixed initial point $M_0(x_0, y_0)$ and variable terminal point $M(x, y)$:

$$I(x, y) = \int_{(x_0, y_0)}^{(x, y)} P(x, y) \, dx + Q(x, y) \, dy \quad (")$$

Let the points M_0 and M , together with the contour of integration, belong to a simply connected domain where all the conditions enumerated at the beginning of the present section are fulfilled. The value of this integral depends solely on the coordinates of the terminal point $M(x, y)$ and thus is a function of x and y .

Under the conditions stated the element of integration $P(x, y) \, dx + Q(x, y) \, dy$ is a total differential, and there is an infinitude of functions of two variables for which the given expression is the total differential; all these functions are described by the formula $u(x, y) + C$ where $u(x, y)$ is one of these functions and C is an arbitrary constant (see Sec. 115). Every such function will be called a *primitive* of the total differential $P \, dx + Q \, dy$. According to the theorem II of Sec. 141, the function $I(x, y)$ is also a primitive of $P \, dx + Q \, dy$; hence,

$$I(x, y) = u(x, y) + C$$

The arbitrary constant C is readily found from the condition $I(x_0, y_0) = 0$: this yields $u(x_0, y_0) + C = 0$, and thus $C = -u(x_0, y_0)$. Therefore

$$I(x, y) = \int_{(x_0, y_0)}^{(x, y)} P(x, y) \, dx + Q(x, y) \, dy = u(x, y) - u(x_0, y_0)$$

The formula obtained is an *analogue of the Newton-Leibniz formula for a line integral of a total differential*. This formula can of course be applied to finding line integrals of total differentials but more often it is used for determining the primitive $u(x, y)$

itself. Such a function is determined to within an arbitrary constant summand, and therefore the constant term $u(x_0, y_0)$ is usually included into that arbitrary constant. This results in the formula

$$u(x, y) = \int_{(x_0, y_0)}^{(x, y)} P(x, y) dx + Q(x, y) dy + C$$

for finding the primitive $u(x, y)$ in any simply connected domain.

In the particular case of a domain of rectangular form treated at the beginning of this section the integration can always be taken along line segments parallel to the axes of coordinates (as shown in Fig. 201); in this case we receive the following two equivalent formulas for computing $u(x, y)$:

$$u(x, y) = \begin{cases} \int_{x_0}^x P(x, y_0) dx + \int_{y_0}^y Q(x, y) dy + C \\ \int_{x_0}^x P(x, y) dx + \int_{y_0}^y Q(x_0, y) dy + C \end{cases} \quad (***)$$

If we put $C=0$ in both formulas we obtain the function $u(x, y)$ vanishing at the point (x_0, y_0) .

In practical calculations the initial point (x_0, y_0) should be chosen so that the element of integration becomes as simple as possible.

Formulas (***) were already derived by means of some other techniques in Sec. 115, I.

Example. Let us find the primitive $u(x, y)$ for the total differential

$$(e^y + x) dx + (xe^y - 2y) dy$$

Taking the origin as the initial point we obtain

$$\begin{aligned} u(x, y) &= \int_0^x (1+x) dx + \int_0^y (xe^y - 2y) dy + C = \\ &= \frac{(1+x)^2}{2} \Big|_0^x + (xe^y - y^2) \Big|_0^y + C = \frac{(1+x)^2}{2} + xe^y - x - y^2 + C \end{aligned}$$

Denoting $\frac{1}{2} + C$ by the same letter C we finally receive

$$u(x, y) = \frac{x^2}{2} + xe^y - y^2 + C$$

This example was already investigated in Sec. 115, I; let the reader compare both methods of the solution.

143. Line Integrals over Space Curves. Up till now we dealt with line integrals of functions of two independent variab-

les x and y taken along paths of integration lying in the plane Oxy . Now we proceed to consider line integrals of functions of three variables. From the point of view of physics such an integral can be interpreted as the work of a *spatial* force field in the motion of a material point in that field. We shall consider *stationary* fields; the forces of a field of this kind are time-independent. Such a force field is described by a function $F(x, y, z)$ or, equivalently, is specified by the projections of the force on the coordinate axes, these projections also being functions of x, y and z . This can be written as

$$\mathbf{F} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

Arguing by analogy with Sec. 138 we arrive at the expression¹

$$\int_L P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz \quad (*)$$

of the work of the field along a space curve L connecting an initial point B with a terminal point C . In what follows, abstracting from the physical meaning, we shall understand P, Q and R as arbitrary continuous functions of three variables. The integrals $\int_L P dx$, $\int_L Q dy$ and $\int_L R dz$ are, respectively, termed *line integrals with respect to the coordinates x, y and z* (integral $(*)$ is also called a *line integral of the second type over a space curve*).

Let L be a smooth curve specified by parametric equations

$$x = x(t), \quad y = y(t) \quad \text{and} \quad z = z(t)$$

the value of the parameter t corresponding to the initial point B being t_B and that corresponding to the terminal point C being t_C . Then integral $(*)$ is transformed into an ordinary definite integral with respect to the variable t :

$$\begin{aligned} & \int_L P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz = \\ & = \int_{t_B}^{t_C} \{P[x(t), y(t), z(t)]x'(t) + Q[x(t), y(t), z(t)]y'(t) + \\ & \quad + R[x(t), y(t), z(t)]z'(t)\} dt \end{aligned}$$

Let the reader derive from this general formula the particular formulas for the cases when the curve L is a line segment parallel to one of the coordinate axes.

For the integral over a space curve of type $(*)$ there is also a close connection between the condition of path-independence and the condition that the element of integration is the total diffe-

¹ Let the reader carry out the corresponding calculations.

rential of a function of three arguments (see Sec. 115, II). To state the corresponding theorems we introduce the definition of a *simply connected spatial domain*:

A three-dimensional domain Ω is said to be *simply connected* if for any piecewise smooth (spatial) simple closed contour L belonging to the domain Ω there is a surface "pulled over" that contour (i.e. having this contour as its boundary) lying entirely within the domain Ω .¹ In other words, this means that any closed contour lying inside the domain Ω can be "contracted" to a point without leaving the domain. Examples of simply connected spatial domains are a polyhedron, a ball, the interior of an ellipsoid, etc. A spherical shell (i.e. a domain bounded by two concentric spheres) is also simply connected. There also exist unbounded simply connected spatial domains such as a *half-space* (i.e. an infinite part of space lying on one side of a given plane), the exterior of a sphere, a *plane layer* (i.e. a part of space between two parallel planes) and the like. An example of a not simply connected domain is the interior of a torus (anchor ring; see Fig. 124). Indeed, a torus is generated by the rotation, in space, of a circle about an axis in its plane but not cutting the circle, and the closed path traced by any point of the rotating circle cannot be contracted to a point without leaving the torus.

Theorem. Let functions $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ of the variadles x , y and z be continuous together with their partial derivatives in a simply connected domain Ω . Then the following four assertions are equivalent:

1. The line integral $\int_L Pdx + Qdy + Rdz$ taken over any closed contour entirely lying within the domain Ω is equal to zero.
2. The line integral $\int_L Pdx + Qdy + Rdz$ is independent of the path of integration connecting two given points.
3. The expression $Pdx + Qdy + Rdz$ is a total differential.
4. At all the points of the domain Ω the equalities

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z} \quad (**)$$

take place.

The proof of this theorem is analogous to that of the corresponding results for the case of an integral over a plane curve, and we shall not present it here (this proof is based on a special generalization of Green's theorem known as Stokes' theorem; the latter will be studied in Sec. 148).

¹ Another type of "connectedness" of a spatial domain (the so-called *acyclicity*) will be defined in Sec. 155*.

We remind the reader that, as was shown in Sec. 115, II, equalities (**) in the fourth assertion are sufficient for the expression $P dx + Q dy + R dz$ to be a total differential. But the sufficiency of these conditions and the method for determining a function $u(x, y, z)$ from its total differential were only established for a rectangular domain of a special type. Repeating almost literally the argument of Sec. 141 we can derive the *analogue of the Newton-Leibniz formula*

$$\int_{(x_0, y_0, z_0)}^{(x, y, z)} P dx + Q dy + R dz = u(x, y, z) - u(x_0, y_0, z_0)$$

and the formula for finding the primitive $u(x, y, z)$ in any simply connected domain:

$$u(x, y, z) = \int_{(x_0, y_0, z_0)}^{(x, y, z)} P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz + C$$

where $M_0(x_0, y_0, z_0)$ is a point belonging to the domain Ω and C an arbitrary constant.

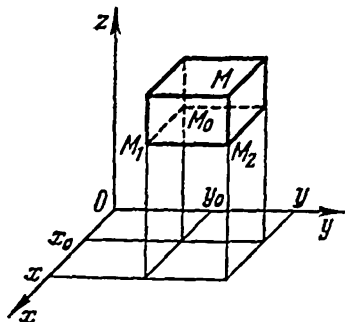


Fig. 203

To compute $u(x, y, z)$ it is most convenient to take as the contour of integration a (spatial) broken line with segments parallel to coordinate axes. If the domain Ω is a parallelepiped of the type considered in Sec. 115, II then, together with any points $M_0(x_0, y_0, z_0)$ and $M(x, y, z)$, it contains the entire rectangular parallelepiped shown in Fig. 203.

For example, let us take as the contour of integration the broken line $M_0M_1M_2M$ (Fig. 203). For M_0M_1 we have $y=y_0$, $z=z_0$ and $dy=0$, $dz=0$; for M_1M_2 we have $z=z_0$, $dx=0$ and $dz=0$; finally, for M_2M we have $dx=0$ and $dy=0$. Hence,

$$u(x, y, z) = \int_{x_0}^x P(x, y_0, z_0) dx + \int_{y_0}^y Q(x, y, z_0) dy + \int_{z_0}^z R(x, y, z) dz + C$$

Analogous formulas are obtained if the integration is performed along other edges of the parallelepiped forming a (spatial) polygonal line joining the point M_0 to the point M ; in Sec. 115, II we presented some other considerations leading to the same formulas.

Example. Let us find the primitive for the total differential

$$du = (2xyz + \ln y) dx + \left(x^2z + \frac{x}{y}\right) dy + (x^2y - 2z) dz$$

In this case conditions (**) are fulfilled for the half-space $y > 0$. As the initial point of integration we can choose, for instance, the point $M_0(0, 1, 0)$; this leads to

$$u(x, y, z) = \int_1^y \frac{x}{y} dy + \int_0^z (y^2 - 2z) dz + C = x \ln y + x^2 yz - z^2 + C$$

Note that the origin $(0, 0, 0)$ cannot be chosen as the initial point since it does not belong to the domain in question (the functions P and Q are not defined at the origin). In Sec. 115, II we applied another technique to solve this problem.

144. Application of Line Integrals to Problems of Mechanics and Thermodynamics.

I. Let us come back to the problem of computing the work of a field of force specified by an equation

$$\mathbf{F} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

If the projections P, Q, R of the force $\mathbf{F}$ on the coordinate axes satisfy the conditions

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z} \quad (*)$$

and the field is defined in a simply connected domain the work of this field is expressed by the line integral

$$\int_L P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz$$

independent of the path of integration. Such a force field is called *potential*. The function $u(x, y, z)$ for which the expression $Pdx + Qdy + Rdz$ is its total differential is called the *potential*¹ of that field of force. Thus, in accordance with the results of Sec. 143, we can say that *the work in a potential field of force is equal to the potential difference between the initial and terminal points*.

Let us consider two simple examples.

(1) *Field of gravitation.* Let a material point of mass m move in a vertical plane under the action of gravitation. Take in this plane a Cartesian coordinate system with the axis Oy directed vertically downward (Fig. 204). Then the projections on Ox and Oy of the force of gravity acting upon the mass point are, respectively,

$$P(x, y) = 0 \quad \text{and} \quad Q(x, y) = mg$$

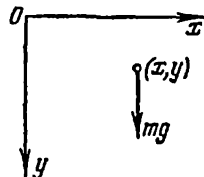


Fig. 204

¹ The potential is sometimes defined as $-u(x, y, z)$. To distinguish between these approaches the term *potential function* is sometimes applied to the function $u(x, y, z)$ but this terminology is not commonly used, and therefore in the translation of this book both terms are used synonymously.—Tr.

where g is the acceleration of gravity. The functions P and Q being constant, conditions (*) are obviously fulfilled (the right-hand and the left-hand members both turn into zero). Consequently, the work performed in the motion of the material point from a point $M_0(x_0, y_0)$ to a point $M(x, y)$ is given by the integral

$$A = \int_{(x_0, y_0)}^{(x, y)} mg \, dy = mg \int_{y_0}^y dy = mgy = mgy_0$$

The work is positive if $y > y_0$, i.e. if the point moves downward, and is negative if it moves upward. The potential function $u(x, y)$ has the form

$$u(x, y) = mgy + C$$

The arbitrary constant C is usually chosen so that $u(x, y) = 0$ for $y = 0$. Then the potential of the force of gravity is the function mgy .

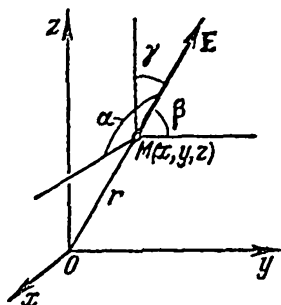


Fig. 205

(2) Electrostatic field generated by a point charge. Let an electric charge $+q$ be placed at the origin. Then the force E with which it acts upon the positive unit charge placed at a point $M(x, y, z)$ is determined by Coulomb's¹ law:

$$|E| = \frac{q}{r^2}$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the distance from the point M to the origin (here Coulomb's law is written for vacuum in the system of units CGSE). The direction of the force E coincides with that of the radius vector $\overline{OM}$. Therefore (see Fig. 205), we have

$$P = |E| \cdot \cos \alpha = \frac{q}{r^2} \frac{x}{r} = \frac{qx}{r^3}$$

$$Q = |E| \cdot \cos \beta = \frac{q}{r^2} \frac{y}{r} = \frac{qy}{r^3}$$

$$R = |E| \cdot \cos \gamma = \frac{q}{r^2} \frac{z}{r} = \frac{qz}{r^3}$$

Now, taking into account that

$$\frac{\partial r}{\partial x} = \frac{x}{\sqrt{x^2 + y^2 + z^2}} = \frac{x}{r}, \quad \frac{\partial r}{\partial y} = \frac{y}{r}, \quad \frac{\partial r}{\partial z} = \frac{z}{r}$$

we easily find the derivatives of the functions P , Q and R by

¹ Coulomb, C. A. (1736-1806), a French physicist.

applying the differentiation rule for a composite function:

$$\frac{\partial P}{\partial y} = -\frac{3qx}{r^4} \frac{\partial r}{\partial y} = -\frac{3qxy}{r^6}, \quad \frac{\partial Q}{\partial x} = -\frac{3qy}{r^4} \frac{\partial r}{\partial x} = -\frac{3qxy}{r^6}$$

etc.

We see that $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$. The reader can readily check that the other two conditions (*) are also fulfilled. These three conditions hold throughout the whole space except at the origin; the space with the point $O(0, 0)$ deleted is an unbounded simply connected domain, and thus the field in question is potential in that domain. Let us find its potential function $u(x, y, z)$. To avoid lengthy calculations note that

$$\begin{aligned} P(x, y, z)dx + Q(x, y, z)dy + R(x, y, z)dz &= \\ = q \frac{x dx + y dy + z dz}{r^3} &= q \frac{\frac{\partial r}{\partial x} dx + \frac{\partial r}{\partial y} dy + \frac{\partial r}{\partial z} dz}{r^3} = q \frac{dr}{r^3} = d\left(-\frac{q}{r}\right) \end{aligned}$$

which means that the function $-\frac{q}{r} + C$ serves as the sought-for potential. Choosing C so that the potential vanishes at infinity we obtain

$$u(x, y, z) = -\frac{q}{r} = -\frac{q}{\sqrt{x^2 + y^2 + z^2}}$$

Hence, the work of the forces of the electrostatic field performed in the motion of unit charge from the point $M_0(x_0, y_0, z_0)$ to the point $M(x, y, z)$ is given by the formula

$$A = -\frac{q}{r} + \frac{q}{r_0}$$

where r_0 and r are the lengths of the radius vectors of the points M_0 and M . Both charges being positive, the work is positive as the distance between them increases ($r > r_0$) and negative as it decreases ($r < r_0$).

II. Let us proceed to apply line integrals to some problems of *thermodynamics*.

In physics the state of a substance is understood as a collection of physical magnitudes completely specifying its physical properties. In thermodynamics these magnitudes are usually the *pressure* p , the *volume* v and the *absolute temperature* T . Hence, in thermodynamics the state is specified by the values of these three magnitudes: p , v and T . But they are connected by an equation (the so-called *equation of state*), and therefore the state is in fact specified by two quantities, for instance, by p and v , the third magnitude T being their function.

Every state can be represented geometrically by a point $M(p, v)$ in the plane Opv , and therefore every process in which the state

of the substance changes is described by a curve in this plane termed the *thermodynamical diagram of the process*. If the thermodynamical system returns to its initial state the process is termed a *cycle*; its diagram is a closed curve.

Let us consider an *ideal (perfect) gas*, that is, a gas obeying the Mendelev-Clapeyron equation of state:

$$pv = RT \quad (R = \text{const})$$

We pose the problem of determining the amount of heat Q absorbed (or rejected) by gas in the process represented by a given diagram L . Consider an arbitrary infinitesimal portion of the curve L from a point $M(p, v)$ to a point $M_1(p+dp, v+dv)$ describing "an infinitesimal" process, the amount of heat corresponding to this process being denoted ΔQ . By the first law of thermodynamics, this amount of heat is expended on changing the internal energy of the particles of gas, that is, on changing the temperature by an amount dT , and on the work performed as the volume gains an increment dv . We can find the element dQ (the principal part of ΔQ linear with respect to dT and dv) assuming that the whole amount of heat transmitted to gas (or rejected by it) is the sum of the following two amounts: (1) an amount expended on changing the temperature by dT for constant volume v and (2) an amount expended on the expansion dv of the gas for constant temperature T . The former is equal to $c_v dT$ where c_v is the *heat capacity in an isochoric process* (i.e. for $v = \text{const}$) of the volume v of gas and the latter is equal to $p dv$.

Consequently, for the corresponding choice of unit measures, we obtain

$$d'Q = c_v dT + p dv$$

where the prime in the symbol $d'Q$ indicates the fact that, as will be seen, this quantity is not a total differential.

In accordance with the Mendelev-Clapeyron equation we can write

$$dT = \frac{v}{R} dp + \frac{p}{R} dv$$

On substituting this expression into the formula of $d'Q$ we obtain

$$d'Q = \frac{c_v}{R} v dp + \frac{c_v + R}{R} p dv$$

Let us discuss the physical meaning of the expression $c_v + R$. For constant pressure p ($dp = 0$) we find from the last two formulas that

$$dT = \frac{p}{R} dv \quad \text{and} \quad d'Q = (c_v + R) \frac{p}{R} dv$$

that is,

$$d'Q = (c_v + R) dT$$

whence it follows that the coefficient in dT is nothing but the *heat capacity in an isobaric process* (i.e. for $p = \text{const}$):

$$c_v + R = c_p \quad (*)$$

(c_p and c_v are considered *constant*).

Thus, ultimately,

$$d'Q = \frac{c_v}{R} v dp + \frac{c_p}{R} p dv$$

To find the desired amount of heat Q it only remains to integrate the expression of $d'Q$ along the diagram of the process L :

$$Q = \int_L \frac{c_v}{R} v dp + \frac{c_p}{R} p dv$$

In this case the conditions for path-independence of the integral are violated since

$$\frac{\partial}{\partial v} \left(\frac{c_v}{R} v \right) = \frac{c_v}{R}, \quad \frac{\partial}{\partial p} \left(\frac{c_p}{R} p \right) = \frac{c_p}{R}$$

and, as follows from relation (*), $c_p \neq c_v$. Thus, the magnitude Q essentially depends on the contour L , that is, Q is not a (single-valued) function of p and v . Our argument shows that the amount of heat absorbed (or rejected) by gas is not a single-valued function of the state and depends not only on the terminal state but also on the course of the process, or, in other words, on all the intermediate states leading to that terminal state. In particular, there may be absorption or rejection of heat in a cycle.

There is a special magnitude, the so-called *entropy*, introduced in thermodynamics (denoted S) which is defined as the integral

$$S = \int_L \frac{d'Q}{T}$$

where L is the diagram of the process.

For an ideal gas

$$S = \int_L \frac{d'Q}{T} = \int_L \frac{c_v}{RT} v dp + \frac{c_p}{RT} p dv$$

and, since $RT = pv$, we obtain

$$S = \int_L \frac{c_v}{p} dp + \frac{c_p}{v} dv$$

The element of integration in the latter integral is a total differential because

$$\frac{\partial}{\partial v} \left(\frac{c_v}{p} \right) = \frac{\partial}{\partial p} \left(\frac{c_p}{v} \right) = 0$$

Thus, the entropy is a (single-valued) function of the state of the gas; its magnitude is independent of the course of the process leading from a given initial state to a given terminal state.

On integrating, we receive

$$S = \ln C p^{c_v} v^{c_p}$$

If the process is *adiabatic* (this means that $Q = \text{const}$, $d'Q = 0$ and, hence, $dS = 0$ and $S = \text{const}$) then

$$p^{c_v} v^{c_p} = \text{const}$$

or, equivalently,

$$p v^k = \text{const}$$

where $k = \frac{c_p}{c_v} > 1$. The latter relation is the equation of the diagram of an adiabatic process (the so-called *adiabatic curve*) for an ideal gas.

145. Line Integrals with Respect to Arc Length. Let $f(x, y)$ be a continuous function in a plane domain D and let L be a smooth curve lying entirely in this domain. Break up the curve L into n parts, choose an arbitrary point (ξ_k, η_k) in each part and form the following *integral sum*:

$$\sum_{k=1}^n f(\xi_k, \eta_k) \Delta s_k \quad (*)$$

where Δs_k is the length of the k th part of the line L . The distinction between this integral sum and that leading to the line integrals with respect to coordinates is that in the former the values of the function are multiplied by the *lengths* of the arcs into which the curve is divided while in the latter they are multiplied by the *projections* of these arcs on the corresponding coordinate axes.

Definition. The limit of integral sum (*), as $n \rightarrow \infty$ and the length of the longest of the arcs into which the contour L is divided tends to zero, is called a *line integral with respect to arc length* (provided that the limit exists).

(As was mentioned in Sec. 138, such an integral is also termed a *line integral of the first type*.)

If this limit exists we write

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n f(\xi_k, \eta_k) \Delta s_k = \int_L f(x, y) ds$$

The properties of the line integral with respect to arc length are completely analogous to those of the definite integral, and we shall not enumerate them here. The computation of a line integral with respect to arc length reduces to the computation of an ordinary definite integral. The rules for this reduction being analogous to those stated in Sec. 139, we confine ourselves to stating the final result:

The Computation Rule for a Line Integral with Respect to Arc Length. To transform a line integral

$$\int_L f(x, y) ds$$

over a curve L specified by parametric equations $x=x(t)$, $y=y(t)$ to an ordinary definite integral one should put

$$x=x(t), \quad y=y(t), \quad ds=\sqrt{x'^2(t)+y'^2(t)} dt$$

in the element of integration and take the definite integral over the interval of variation of t corresponding to the given path of integration.

This general transformation rule for a line integral with respect to arc length implies the corresponding rules for the particular cases when the equation of the contour of integration is written as $y=y(x)$ or $x=x(y)$. In the former case we should put $t=x$ in the general formula and in the latter $t=y$.

If L is a piecewise smooth curve it should be broken up into smooth parts, and the integral over the whole contours should be computed as the sum of the integrals over these parts.

Example. Let us evaluate the integral

$$\int_L y ds$$

where L is the arc of the parabola $y^2=2x$ connecting its points $(0, 0)$ and $(4, \sqrt{8})$. In this case it is convenient to represent the path of integration in the form of an equation solved with respect to x : $x=\frac{y^2}{2}$. Then $x'=y$, and the given integral goes into

$$\int_L y ds = \int_0^{\sqrt{8}} y \sqrt{1+y^2} dy = \frac{(1+y^2)^{\frac{3}{2}}}{3} \Big|_0^{\sqrt{8}} = \frac{26}{3}$$

Line integrals with respect to arc length are particularly convenient for solving the problem of determining the mass of a material line L (see Sec. 86, III) if the linear density δ is specified not as a function of the arc length of the line but as a function of the coordinates x and y of the moving point M of the line:

$\delta = \delta(x, y)$. Then the mass of the line is equal to $\int_L \delta(x, y) ds$ (let the reader justify this formula).

Let us discuss a geometrical problem which is also readily solved with the aid of the line integral with respect to arc length. Let there be given a cylindrical surface G whose directrix is a line L in the plane Oxy and whose elements are perpendicular to that plane (Fig. 206). The problem is to compute the area Q of

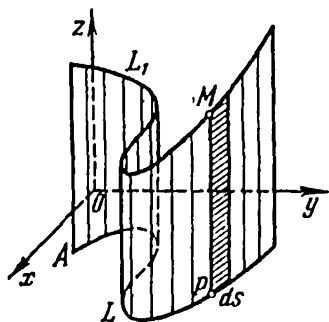


Fig. 206

the part of the surface between the line L and a given curve L_1 on the same surface lying above L . The z -coordinate of the moving point M of the curve L_1 is a function of the coordinates of the variable point P of the line L onto which it is projected (see Fig. 206):

$$z = f(x, y)$$

Let us take an infinitesimal element of arc length ds at an arbitrary point P of the line L . The part of the surface in question corresponding to ds is shaded in Fig. 206. The principal part (proportional to ds) of the area of this surface element is the area of the rectangle with base ds and altitude equal to the z -coordinate of the corresponding point M of the curve L_1 . Hence, the differential dQ of the sought-for area is

$$dQ = f(x, y) ds$$

The integration of this expression over the line L results in

$$Q = \int_L f(x, y) ds$$

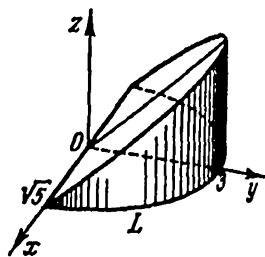


Fig. 207

Example. Let us determine the area Q of the lateral surface of the half of the elliptical cylinder $\frac{x^2}{5} + \frac{y^2}{9} = 1, y \geq 0, z \geq 0$, truncated by the plane $z = y$ (see Fig. 207).

We have

$$Q = \int_L z ds = \int_L y ds$$

where L is the arc of the ellipse $\frac{x^2}{5} + \frac{y^2}{9} = 1, y \geq 0$.

Let us evaluate the integral. The contour of integration is represented parametrically as $x = \sqrt{5} \cos t, y = 3 \sin t$, and, conse-

quently,

$$Q = \int_0^{\pi} 3 \sin t \sqrt{9 \cos^2 t + 5 \sin^2 t} dt = 3 \int_0^{\pi} \sin t \sqrt{4 \cos^2 t + 5} dt$$

Putting $\cos t = u$ we find

$$Q = 3 \int_{-1}^1 \sqrt{4u^2 + 5} du = 6 \int_0^1 \sqrt{4u^2 + 5} du$$

Now, using Integral 34 (with $a=2$, $b=\sqrt{5}$) in the table of integrals at the end of the book we finally receive

$$Q = 3 \left[u \sqrt{4u^2 + 5} + \frac{5}{2} \ln |2u + \sqrt{4u^2 + 5}| \right]_0^1 = 3 \left(3 + \frac{5}{4} \ln 5 \right)$$

The definition of the line integral with respect to arc length

$$\int_L f(x, y, z) ds$$

taken along a space curve is analogous to that stated for the case of a plane curve, and we shall not dwell on it here.

It should be noted that the line integral

$$\int_L P dx + Q dy + R dz$$

considered in Sec. 143 can always be written as an integral with respect to arc length. Indeed, according to the formulas given at the end of Sec. 68, we have

$$dx = ds \cos \alpha, \quad dy = ds \cos \beta, \quad dz = ds \cos \gamma$$

where $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ are the direction cosines of the tangent vector $\mathbf{T}$ to the curve L and ds is the differential of arc length; the direction of the vector $\mathbf{T}$ is specified by the chosen direction of integration along the line L . On replacing dx , dy and dz by their expressions we bring the given integral to the form

$$\int_L P dx + Q dy + R dz = \int_L (P \cos \alpha + Q \cos \beta + R \cos \gamma) ds$$

§ 2 *. Surface Integrals

146*. Fluid Flux Across a Surface. The Surface Integral.

We shall begin the study of surface integrals with considering a concrete physical problem. Suppose that there is a flow of fluid. The flow is said to be *steady-state* or *stationary* if the velocity of the particles of fluid flowing through any given point depends

solely on that point and is independent of time. Consequently, at each point $M(x, y, z)$ of the domain occupied by a steady-state fluid flow a velocity vector $\mathbf{v}(x, y, z)$ of the particles of fluid is specified or, in other words, a *vector field* of velocities of the fluid¹ is defined in this domain. The projections of the vector $\mathbf{v}(x, y, z)$ on the coordinate axes are functions of the coordinates of the point M , which can be written as

$$\mathbf{v} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

We shall assume that the density of the fluid is constant and, conditionally, equal to unity. Now we pose the problem of determining the amount of fluid flowing in unit time through a given surface; it is of course assumed that the particles of the fluid can pass freely through that surface. This amount of fluid is called the *flux of the fluid across the given surface*.

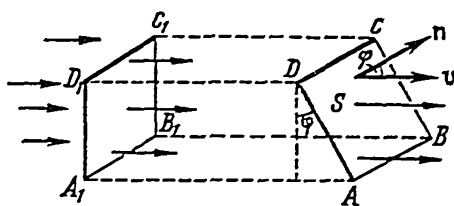


Fig. 208

Let us begin with the simplest case. Suppose that the velocity of the flow is the same at all the points: $\mathbf{v} = \text{const}$. The flux of the fluid through the rectangle $A_1B_1C_1D_1$ (see Fig. 208) placed in a plane perpendicular to the velocity $\mathbf{v}$ is equal to the product of the area of the rectangle by the magnitude of the velocity; hence, denoting this flux by K , we can write

$$K = S_{A_1B_1C_1D_1} |\mathbf{v}|$$

It is apparent that the same amount K of fluid passes through the rectangle $ABCD$ lying in a plane whose normal vector $\mathbf{n}$ forms an angle φ with the velocity $\mathbf{v}$ (see Fig. 208). If the area of the rectangle $ABCD$ is denoted S the relations $D_1C_1 = DC$ and $A_1D_1 = AD \cos \varphi$ imply that $S_{A_1B_1C_1D_1} = S \cos \varphi$. Therefore we have

$$K = S |\mathbf{v}| \cos \varphi = S v_n$$

where v_n is the projection of the velocity $\mathbf{v}$ on the normal $\mathbf{n}$. This formula applies not only to a rectangle but also to the flux K of the fluid across an arbitrary plane area whose normal forms the same angle φ with the velocity $\mathbf{v}$ (see Fig. 209). If the area of such a plane figure is σ the area of its projection on a plane

¹ The velocity field of a fluid flow is a concrete example of a general notion of a *vector field*; another example of a vector field is a *field of force* which was dealt with in Sec. 138 in connection with the problem of computing the work performed in the motion of a material point. The general theory of vector fields will be considered in § 3 of this chapter.

perpendicular to the velocity $\mathbf{v}$ is equal to $\sigma \cos \varphi$.¹ If a plane area is perpendicular to the velocity of the flow the flux is equal to the product of the area by the magnitude of the velocity.

Now we pass to the general case. Let there be a domain in which a field of velocities of a fluid flow is specified by a vector function

$$\mathbf{v} = P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$$

Take an arbitrary surface S . To determine the flux across this surface let us break up the given surface into n parts of arbitrary shape and denote by $\Delta\sigma_1, \Delta\sigma_2, \dots, \Delta\sigma_n$ the areas of these parts. Next we choose

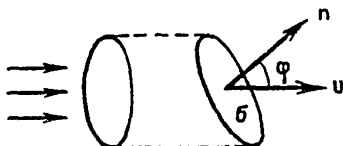


Fig. 209

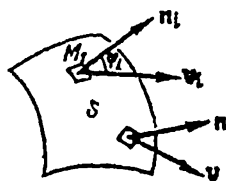


Fig. 210

an arbitrary point $M_i(x_i, y_i, z_i)$ on each part (Fig. 210). Considering (approximately) each part as being plane and the velocity of the fluid as being constant and equal to its value at the point M_i within each part, we can write down an *approximation* to the fluid flow across the whole area:

$$\begin{aligned} K_n &= \Delta\sigma_1 |\mathbf{v}_1| \cos \varphi_1 + \Delta\sigma_2 |\mathbf{v}_2| \cos \varphi_2 + \dots + \Delta\sigma_n |\mathbf{v}_n| \cos \varphi_n = \\ &= \sum_{i=1}^n \Delta\sigma_i |\mathbf{v}_i| \cos \varphi_i \end{aligned}$$

Here $\mathbf{v}_i$ is the (vectorial) value of the velocity at the point M_i and φ_i is the angle between the normal $\mathbf{n}_i$ to the surface S at the point M_i and the velocity vector $\mathbf{v}_i$.

Noting that $|\mathbf{v}_i| \cos \varphi_i$ is the *scalar product* of the vector $\mathbf{v}_i$ by the unit normal vector we shall transform this approximate expression. Let α_i, β_i and γ_i be the angles formed by the normal $\mathbf{n}_i$ with the coordinate axes (these angles depend of course on the point M_i). Then the unit normal vector at the point M_i has the direction cosines $\cos \alpha_i, \cos \beta_i$ and $\cos \gamma_i$ as its projections. Consequently,

$$|\mathbf{v}_i| \cos \varphi_i = P_i \cos \alpha_i + Q_i \cos \beta_i + R_i \cos \gamma_i$$

where P_i, Q_i and R_i are the values of the corresponding func-

¹ The whole area can be thought of as being divided into a great number of small rectangles; since the formula is valid for every such rectangle it also applies to the whole area.

tions P , Q and R at the point M_i . Therefore

$$K_n = \sum_{i=1}^n (P_i \cos \alpha_i + Q_i \cos \beta_i + R_i \cos \gamma_i) \Delta \sigma_i \quad (*)$$

Now, passing to the limit as $n \rightarrow \infty$ on condition that every elementary area $\Delta \sigma_i$ contracts to a point we obtain the true value of the fluid flux K :

$$K = \lim K_n = \lim \sum_{i=1}^n (P_i \cos \alpha_i + Q_i \cos \beta_i + R_i \cos \gamma_i) \Delta \sigma_i$$

This limit is written as

$$\iint_S (P \cos \alpha + Q \cos \beta + R \cos \gamma) d\sigma$$

and called the *surface integral* (over the surface S). The expressions P , Q and R and the direction cosines $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ in the element of integration are functions of the coordinates of the variable point $M(x, y, z)$ of the surface S . In the particular case when S is a part of a plane the direction cosines are constant.

For the sake of brevity, we shall use the notation

$$P \cos \alpha + Q \cos \beta + R \cos \gamma = f(M)$$

where $f(M)$ symbolizes a function of the point M on the surface S .

Now we proceed to the general definition of a surface integral abstracting from the concrete conditions of the problem. Let S be a smooth surface and $f(M)$ a continuous function of the variable point M of the surface. The function $f(M)$ can be written as $f(x, y, z)$ where x , y and z are the coordinates of the point M . Dividing the surface S into n parts as was done above we form the *integral sum*

$$\sum_{i=1}^n f(M_i) \Delta \sigma_i$$

Definition. The limit of the integral sum as $n \rightarrow \infty$ and the diameter of each elementary part tends to zero is called an *integral over the surface S* (provided that the limit exists).

If this limit exists we write

$$\lim \sum_{i=1}^n f(M_i) \Delta \sigma_i = \iint_S f(M) d\sigma = \iint_S f(x, y, z) d\sigma$$

If, for instance, the function $f(x, y, z)$ describes the surface density of a mass distributed over the surface the integral is equal to the total mass of the material surface; if $f(x, y, z)$ is a density of electric charge distribution the integral expresses the total charge

carried by the surface. In the particular case when $f(x, y, z) \equiv 1$ the integral $\iint_S d\sigma$ is simply equal to the area of the surface S .¹

147*. Properties of Surface Integrals. The general idea of the definition of the surface integral is quite analogous to that of the definition of the double integral. Therefore the assertions concerning the properties of double integrals enumerated in Sec. 129 are extended without any essential changes to surface integrals of the form $\iint_S f(M) d\sigma$ where $f(M)$ is an arbitrary function defined on the surface S . But when the function $f(M)$ is originally specified as $f(M) = P \cos \alpha + Q \cos \beta + R \cos \gamma$ where P , Q and R are given functions defined on S and α , β and γ are the angles formed by the normal to the surface S with the coordinate axes there appear some peculiarities which we are going to discuss here.

Integrals of the form

$$\iint_S (P \cos \alpha + Q \cos \beta + R \cos \gamma) d\sigma \quad (*)$$

are connected with the notion of the flux of the vector field $\mathbf{P} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ across the surface S (this will be discussed in Sec. 152) whose special case is the fluid flux considered in Sec. 146, and its peculiarities are connected with some specific features of the physical problem leading to (*).

The matter is that the normal to a surface has *two* possible directions (see Fig. 211). When we considered an approximation to the fluid flux through a given surface (Sec. 146) we tacitly assumed that all the normals $\mathbf{n}_i$ were directed so that their tips were on *one side*

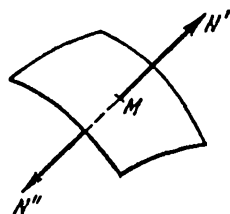


Fig. 211

of the surface S . In what follows we shall distinguish between the two sides of a surface S ; this can be, for instance, thought of as one side of the surface S being painted red and the other side painted blue. Accordingly, we shall distinguish between the *two opposite* directions of the normal (shown as MN' and MN'' in Fig. 211). Consequently, when speaking of an integral of type (*) we should distinguish between the two integrals depending on the *definite side* of the surface over which the integral is taken (i.e. over the "red" or "blue" side). It is clear

¹ Here we proceed from an intuitive concept of area of a surface whose rigorous definition involves some sophisticated considerations and is dealt with in comprehensive courses of mathematical analysis. We do not state this definition referring the reader to such courses. The formula for computing the area of a surface specified by an equation $z = f(x, y)$ will be given in Sec. 148.

that if the direction of the normal is changed to the opposite the direction cosines $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ change sign, which results in the change of the sign of the integral. Thus, integrals (*) taken over opposite sides of the surface S have different signs. If the element of integration $(P \cos \alpha + Q \cos \beta + R \cos \gamma) d\sigma$ is, for brevity, again denoted by $f(M)$, the latter remark is expressed by the relation

$$\iint_{-S} f(M) d\sigma = - \iint_{+S} f(M) d\sigma \quad (**)$$

where $+S$ and $-S$ designate (conditionally) the two sides of the surface S .

We stress that there are physical problems of two types leading to surface integrals $\iint_S f(M) d\sigma$. In problems of one type the integrand $f(M)$ is not connected with the direction of the normal to the surface S , and therefore the corresponding integrals do not depend on the choice of the side of the surface. Such are, for instance, the problems of finding the mass or the charge distributed over a surface with given density δ . In problems of the other type the function $f(M)$ is of the form $P \cos \alpha + Q \cos \beta + R \cos \gamma$ and essentially depends on the direction of the normal; for instance, this is the case in the problem of determining the fluid flux; it involves the integrand $f(M) = P \cos \alpha + Q \cos \beta + R \cos \gamma$ which is a linear function of the direction cosines of the normal. Therefore in the latter problems we must stipulate the side of the surface over which the integration is extended since the change of the side results in the change of sign of the integral¹.

In conclusion let us discuss again the physical meaning of the sign of surface integral (*) in connection with the computation of fluid flux across a surface S . Suppose that a definite direction of the normal to S is chosen. On different parts of the surface S the fluid can flow in different directions, that is, along the chosen normal or in the opposite direction. If the resultant flux K is positive the total amount of fluid flowing in the chosen direction exceeds that in the opposite direction, and if the flux is negative the former is less than the latter.

Remark. In our discussion we meant that the surfaces in question were such that it was possible to distinguish between their two opposite sides. Surfaces of this kind are termed *two-sided*.

¹ The reader should note that a similar situation appears in the study of line integrals. For instance, the problem of finding the work of a force field is essentially connected with the direction of the tangent line to the contour of integration (drawn in the direction of motion) while the problem of computing the mass of a material line or of the area of a cylindrical surface (Sec. 145) is not connected with it.

If L is the boundary of a two-sided surface it is only possible to pass from one side to the other when L is intersected by the path of motion. But there also exist *one-sided surfaces*. An example of this kind is the so-called *Möbius strip* shown in Fig. 212. It can be obtained by taking a rectangular strip of paper and pasting its two opposite ends together after giving it half a twist (if they are pasted together without twist we obtain an ordinary cylindrical surface which is two-sided). It is easily seen that after we traverse the centre line of the Möbius strip the direction of the normal to the surface is reversed, which means that the surface is in fact one-sided. Hence, the Möbius strip cannot be painted so that its one side is red and the other blue! In what follows we shall always deal with two-sided surfaces.

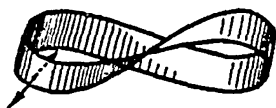


Fig. 212

If we have an integral (*) taken over a definite side of a surface S it satisfies condition (**). Hence, the theorem on splitting the surface of integration into parts should be stated for such integrals as follows:

If the surface S is divided into parts $S_1, S_2, \dots, S_k$ and the directions of the normals to these parts are chosen so that they are coherent with the direction chosen for the entire surface S the integral over the entire surface is equal to the sum of the integrals taken over the parts:

$$\iint_S f(M) d\sigma = \iint_{S_1} f(M) d\sigma + \dots + \iint_{S_k} f(M) d\sigma$$

We shall often encounter surface integrals of type (*) taken over *closed* surfaces. For an integral over a closed surface we should distinguish between the *outer* and *inner* sides of the surface of integration. The outer side is usually called the *positive side* and the inner side is called *negative*. For integrals (*) over closed surfaces we can state the following theorem analogous to the theorem on line integrals over closed contours (see pp. 490, 491):

Theorem. Let Ω be a spatial domain bounded by a closed surface S . If Ω is broken up into n parts with the aid of closed surfaces $S_1, S_2, \dots, S_k$ the integral taken over the whole surface S (for instance, over its positive side) is equal to the sum of the integrals taken over the corresponding (i.e. positive) sides

of the surfaces $S_1, S_2, \dots, S_k$:

$$\iint_S f(M) d\sigma = \iint_{S_1} f(M) d\sigma + \iint_{S_2} f(M) d\sigma + \dots + \iint_{S_k} f(M) d\sigma$$

Proof. The right-hand side of this equality involves the integrations over both sides of each portion of the surfaces $S_1, S_2, \dots, S_k$ lying inside the domain Ω , and the sum of the corresponding integrals is equal to zero. Therefore there only remain the integrals over the parts of the boundary of the domain Ω whose collection forms the entire surface S . Hence, the sum on the right is equal to the surface integral over the whole surface S , which is what we intended to prove.

148*. Evaluating Surface Integrals. Consider a surface integral

$$\iint_S f(x, y, z) d\sigma$$

over a surface S specified by an equation $z = z(x, y)$. It is supposed that the function z and the derivatives z'_x and z'_y are continuous (such a surface is said to be *smooth*).

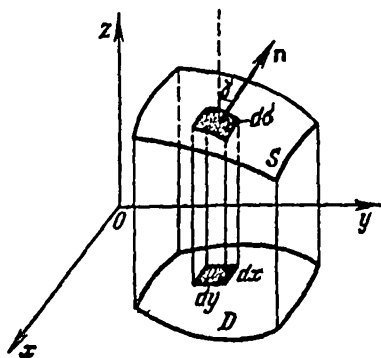


Fig. 213

We shall show that in this case the computation of the surface integral over the surface S is reduced to the computation of a double integral over the domain D serving as the projection of the surface S on the plane Oxy (Fig. 213).

Let us take an element of area $dx dy$ of the domain D and consider the corresponding element of area $d\sigma$ of the surface S projected on the former. Next we draw the normal n to the element $d\sigma$ so that the angle γ between this normal and the axis Oz is acute. We shall assume, approximately, that the direction of the normal n remains invariable within the area $d\sigma$, which means that the surface element $d\sigma$ is approximated by a portion of the tangent plane to the surface drawn through the point of application of the normal. Then, as was indicated in Sec. 146, we have

$$\cos \gamma d\sigma = dx dy$$

According to the results of Sec. 119, the normal to a surface specified by an equation

$$z = z(x, y)$$

has the projections z'_x , z'_y and -1 . Therefore, the cosine of the acute angle γ between the normal and Oz is expressed by the formula

$$\cos \gamma = \frac{1}{\sqrt{1+z_x'^2+z_y'^2}}$$

Thus, we have $d\sigma = \sqrt{1+z_x'^2+z_y'^2} dx dy$, and

$$f(x, y, z) d\sigma = f[x, y, z(x, y)] \sqrt{1+z_x'^2+z_y'^2} dx dy$$

It follows from the latter equality that the "summation" of the expressions on the left, that is, the integration over the surface S , is equivalent to the "summation" of the expressions on the right, that is, to the integration over the domain D :

$$\iint_S f(x, y, z) d\sigma = \iint_D f[x, y, z(x, y)] \sqrt{1+z_x'^2+z_y'^2} dx dy \quad (*)$$

It may turn out that the given surface cannot be represented by an equation solved with respect to z but can be specified by an equation of the form $y = y(x, z)$ or $x = x(y, z)$. Then, respectively, we can apply the same argument to the projections of the surface S on the plane Oxz or Oyz . If the acute angles formed by the normal to the surface with the axes Oy and Ox are denoted β and α we can write the formulas

$$\cos \beta d\sigma = dx dz \text{ and } \cos \alpha d\sigma = dy dz$$

Let the reader derive the corresponding reduction formulas for the integral over the surface to the corresponding double integrals in the case of the equations $y = y(x, z)$ and $x = x(y, z)$.

If the entire surface S cannot be represented by one equation expressing one of the coordinates in terms of the others we should divide the surface into parts for which such a specification is possible (as a rule, a surface admits of such a division) and find the desired integral as the sum of the integrals over the parts.

If the function $f(x, y, z)$ in equality (*) is identically equal to unity we obtain the computation formula for the area of the given surface; denoting this area by the same letter S (used earlier for designating the surface itself) we can write

$$S = \iint_D \sqrt{1+z_x'^2+z_y'^2} dx dy$$

In this formula $z = z(x, y)$ is the equation of the surface and D is its projection on the plane Oxy .

Example. Let us determine the area of the portion of the surface of the paraboloid of revolution $2z = x^2 + y^2$ cut out by the cylinder $x^2 + y^2 = R^2$.

In this case we have $z'_x = x$ and $z'_y = y$; therefore

$$S = \iint_D \sqrt{1+x^2+y^2} \, dx \, dy$$

where D is the projection of that portion of the given surface on the plane Oxy ; this projection is the circle of radius R and centre at the origin in the plane Oxy . On passing to polar coordinates we obtain

$$S = \int_0^{2\pi} d\varphi \int_0^R \sqrt{1+r^2} \, r \, dr = \frac{2\pi}{3} [(1+R^2)^{3/2} - 1]$$

Let us proceed to compute an integral of the form

$$\iint_S [P(x, y, z) \cos \alpha + Q(x, y, z) \cos \beta + R(x, y, z) \cos \gamma] \, d\sigma$$

where $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ are the direction cosines of the normal drawn to the side of the surface over which the integration is performed.

The form of the integrand indicates that it is convenient to evaluate this integral as the sum of the three integrals

$$\iint_S P \cos \alpha \, d\sigma, \quad \iint_S Q \cos \beta \, d\sigma \quad \text{and} \quad \iint_S R \cos \gamma \, d\sigma$$

Let us begin with the integral

$$\iint_S R(x, y, z) \cos \gamma \, d\sigma$$

Suppose that the surface S is specified by an equation $z = z(x, y)$, that is, has the form shown in Fig. 213. We shall distinguish between the *upper* side of this surface (i.e. the one to which the normal n in the figure is drawn) and its *lower* side. The angle γ is acute for the upper side and, as was mentioned above, we have $\cos \gamma \, d\sigma = dx \, dy$. Consequently,

$$R(x, y, z) \cos \gamma \, d\sigma = R[x, y, z(x, y)] \, dx \, dy$$

and therefore

$$\iint_S R(x, y, z) \cos \gamma \, d\sigma = \iint_D R[x, y, z(x, y)] \, dx \, dy \quad (**)$$

where D is the projection of the surface S on the plane Oxy .

It is obvious that for the integral over the lower side of the surface we similarly obtain the expression

$$-\iint_D R[x, y, z(x, y)] \, dx \, dy$$

The other two integrals $\iint_S P \cos \alpha \, d\sigma$ and $\iint_S Q \cos \beta \, d\sigma$ are computed analogously. It should only be taken into account that in the formulas

$$\cos \alpha \, d\sigma = dy \, dz \text{ and } \cos \beta \, d\sigma = dx \, dz$$

the angles α and β are understood as *acute* angles between the normal to the surface and, respectively, the axes Ox and Oy . Therefore, when resolving the equation of the surface with respect to x or y and passing to the corresponding equations $x = x(y, z)$ or $y = y(x, z)$ we obtain the formulas

$$\iint_S Q(x, y, z) \cos \beta \, d\sigma = \iint_{D_1} Q[x, y(z, x), z] \, dx \, dz$$

and

$$\iint_S P(x, y, z) \cos \alpha \, d\sigma = \iint_{D_2} P[x(y, z), y, z] \, dy \, dz$$

for the integrals taken over the side of the surface S whose normal is in the direction of Ox or Oy . These formulas are analogous to (**). If it is necessary to take the integral over the opposite side of the surface the signs of the double integrals in the last formulas should be changed to the opposite. The domains D_1 and D_2 in these formulas are, respectively, the projections of the surface S on the planes Oxz and Oyz .

If, when computing an integral of the type $\iint_S P \cos \alpha \, d\sigma$, $\iint_S Q \cos \beta \, d\sigma$ or $\iint_S R \cos \gamma \, d\sigma$, we see that it is impossible to represent the equation of the surface in the required form (solved with respect to the corresponding coordinate) we should split the surface into parts for which such a representation is possible (provided that such a division of the given surface exists) and take the sum of the integrals over these parts. It should also be noted that if the surface S includes a portion of a cylindrical surface whose elements are, for instance, parallel to Oz , the integral $\iint_S R \cos \gamma \, d\sigma$ taken over this portion is equal to zero since the normal is perpendicular to the element of the cylinder, and therefore $\gamma = \frac{\pi}{2}$ and $\cos \gamma = 0$. The right-hand member of formula (**) also turns into zero for such a portion because the projection of that portion of the surface on the plane Oxy is not a plane region but a line having zero area. Hence, formula (**) remains valid in this case. Similar remarks can be made in connection with cylindrical parts of the surface with elements parallel to Ox or Oy .

In accordance with what has been said, a surface integral of the type we are discussing here is often written as

$$\iint_S (P \cos \alpha + Q \cos \beta + R \cos \gamma) d\sigma = \iint_S P dy dz + Q dx dz + R dx dy$$

In comprehensive courses of mathematical analysis the integral on the right is defined with the aid of some special integral sums and is termed a *surface integral of the second type* (an integral $\iint_S f(M) d\sigma$ is termed a *surface integral of the first type*). We shall not consider such integrals separately.

This form of the representation of the integral is convenient since it indicates directly on which coordinate plane the surface S is projected when the integration in each term is performed.

Examples. (1) Let us evaluate the integral

$$\iint_S x dy dz + dx dz + xz^2 dx dy$$

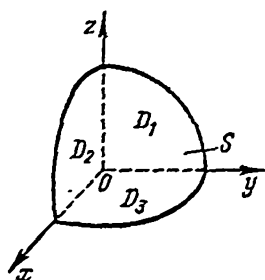


Fig. 214

where S is the outer side of the part of the sphere $x^2 + y^2 + z^2 = 1$ in the first coordinate trihedral.

Let us denote the projections of the surface on the coordinate planes Oyz , Oxz and Oxy as D_1 , D_2 and D_3 , respectively (see Fig. 214). These are quarter circles of radius 1. Then we have

$$I_1 = \iint_S x dy dz = \iint_{D_1} \sqrt{1-y^2-z^2} dy dz$$

$$I_2 = \iint_S dx dz = \iint_{D_2} dx dz$$

and

$$I_3 = \iint_S xz^2 dx dy = \iint_{D_3} x(1-x^2-y^2) dx dy$$

The second integral is simply equal to the area of the domain D_1 , i.e. to $\frac{\pi}{4}$. The first and the third integrals are computed by passing to polar coordinates:

$$I_1 = \int_0^{\frac{\pi}{2}} d\varphi \int_0^1 \sqrt{1-r^2} r dr = \frac{\pi}{2} \left(-\frac{(1-r^2)^{3/2}}{3} \Big|_0^1 \right) = \frac{\pi}{6}$$

and

$$I_2 = \int_0^{\frac{\pi}{2}} \cos \varphi \, d\varphi \int_0^1 r(1-r^2) \, r \, dr = 1 \cdot \left(\frac{r^3}{3} - \frac{r^5}{5} \right) \Big|_0^1 = \frac{2}{15}$$

Thus,

$$I = I_1 + I_2 + I_3 = \frac{\pi}{4} + \frac{\pi}{6} + \frac{2}{15} = \frac{5\pi}{12} + \frac{2}{15}$$

(2) Let us find the integral

$$I = \iint_S z \cos \gamma \, d\sigma$$

where S is the outer side of the sphere $x^2 + y^2 + z^2 = 1$.

In this example the z -coordinate cannot be written as a single-valued function of x and y for the whole surface of integration S . Let us break up the surface into two parts: the part S_1 lying above the plane Oxy and the part S_2 below it. Accordingly, their equations are

$$z = \sqrt{1-x^2-y^2} \quad \text{and} \quad z = -\sqrt{1-x^2-y^2}$$

Hence, we can write

$$I = \iint_S z \cos \gamma \, d\sigma = \iint_{S_1} z \cos \gamma \, d\sigma + \iint_{S_2} z \cos \gamma \, d\sigma$$

Let us pass to the corresponding double integrals. Since S_1 is the outer side of the portion of the sphere above the plane Oxy and thus is its *upper* side, we have

$$\iint_{S_1} z \cos \gamma \, d\sigma = \iint_D \sqrt{1-x^2-y^2} \, dx \, dy$$

where D is the circle $x^2 + y^2 \leq 1$. The surface S_2 is the outer side of the half-sphere lying below the plane Oxy and thus is its *lower* side; consequently,

$$\iint_{S_2} z \cos \gamma \, d\sigma = - \iint_D -\sqrt{1-x^2-y^2} \, dx \, dy$$

where D is the same circle.

The final result is

$$I = \iint_D \sqrt{1-x^2-y^2} \, dx \, dy + \iint_D \sqrt{1-x^2-y^2} \, dx \, dy = 2 \iint_D \sqrt{1-x^2-y^2} \, dx \, dy$$

The integral $\iint_D \sqrt{1-x^2-y^2} \, dx \, dy$ being equal to half the volume of the ball of radius 1, i.e. to $\frac{2}{3}\pi$, we obtain $I = \frac{4}{3}\pi$.

149*. Stokes' ¹ Theorem. There is a formula for surface integrals analogous to Green's formula for line integrals (see Sec. 140) reducing the computation of a surface integral over a surface S bounded by a contour L to a line integral round this contour (Fig. 215). This result is established in the following theorem.

Stokes' Theorem. Let $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ be three arbitrary functions continuous together with their partial derivatives in a three-dimensional domain. Then, for any smooth surface S lying in this domain, we have

$$\iint_S \left[\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \cos \alpha + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \cos \beta + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \cos \gamma \right] d\sigma = \int_L P dx + Q dy + R dz \quad (*)$$

where $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ are the direction cosines of the normal to the surface S and L is its boundary.

Relation (*) expresses *Stokes' formula*.

In this formula the direction of integration round the contour L must be *coherent with the orientation* of the surface S in the sense of the following simple rule: when a person walks along the contour L on the chosen side of the surface S and traverses L in the direction of integration the surface S must always remain on the left (it is supposed that we use a right-handed coordinate system $Oxyz$; see Fig. 215).

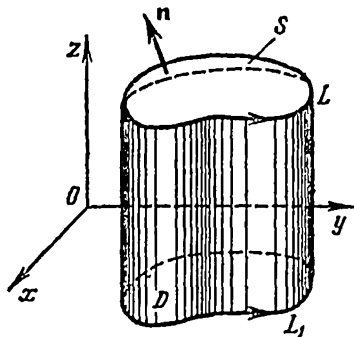


Fig. 215

Proof. The theorem is proved by reducing the surface integral first to the double integral and then, with the aid of Green's formula, to the line integral. We shall begin with the case when the surface S is described

by an equation $z = z(x, y)$. Let us assume that S is the outer side of the surface. Under this condition, $\cos \gamma > 0$, and we can transform the surface integral

$$I = \iint_S \left(\frac{\partial P}{\partial z} \cos \beta - \frac{\partial P}{\partial y} \cos \gamma \right) d\sigma$$

by taking into account that the vector with projections $-z'_x$, $-z'_y$, 1 goes in the direction of the normal to the upper side of the surface. The direction cosines being proportional to the co-

¹ Stokes, G.G. (1819-1903), an English mathematician and physicist.

responding projections, we have

$$\frac{\cos \beta}{\cos \gamma} = \frac{-z'_y}{1} = -z'_y$$

Therefore $\cos \beta = -z'_y \cos \gamma$, and

$$I = - \iint_S \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} z'_y \right) \cos \gamma \, d\sigma$$

Let us reduce the latter surface integral to the double integral. To this end, the variable z in the integrand must be replaced by its expression in terms of x and y in accordance with the equation of the surface $z = z(x, y)$. Then the integrand is transformed to the expression equal to the total partial derivative with respect to y of the composite function resulting from the substitution of $z(x, y)$ for z into the function $P(x, y, z)$. For we have

$$\frac{\partial}{\partial y} P[x, y, z(x, y)] = \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z} z'_y$$

Thus, the integral I is written as

$$I = - \iint_D \frac{\partial P[x, y, z(x, y)]}{\partial y} \, dx \, dy$$

where D is the projection of the surface S on the plane Oxy (see Fig. 215). Now we apply Green's formula (Sec. 140) and pass to the line integral round the boundary L_1 of the domain D :

$$I = \int_{L_1} P[x, y, z(x, y)] \, dx$$

The contour L_1 is the projection of the space curve L (bounding the surface S) on the plane Oxy , and the integration goes in the positive direction of L_1 . To complete our consideration it only remains to use the fact that, since the contour L lies on the surface S , the coordinates of its points satisfy the equation $z = z(x, y)$, and therefore the values of the function $P(x, y, z)$ at the points (x, y, z) of the contour L are equal to the values of the expression $P[x, y, z(x, y)]$ at the points (x, y) of the contour L_1 . The projections of the corresponding parts of the contours L and L_1 (appearing as the contours are partitioned) on the axis Ox obviously coincide, and therefore the integral sums constructed for the line integrals of the function P over the contours L and L_1 with respect to the coordinate x coincide as well. Thus, the integrals are also equal:

$$\int_{L_1} P[x, y, z(x, y)] \, dx = \int_L P(x, y, z) \, dx$$

Hence,

$$I = \iint_S \left(\frac{\partial P}{\partial z} \cos \beta - \frac{\partial P}{\partial y} \cos \gamma \right) d\sigma = \int_L P dx \quad (A)$$

Arguing in the same manner as on page 493 we can show that the resultant formula is in fact applicable to surfaces of a more complicated shape which can be divided into a finite number of parts such that formula (A) applies to each of them.

Using the same techniques we readily derive two more formulas:

$$\iint_S \left(\frac{\partial Q}{\partial x} \cos \gamma - \frac{\partial Q}{\partial z} \cos \alpha \right) d\sigma = \int_L Q dy \quad (B)$$

and

$$\iint_S \left(\frac{\partial R}{\partial y} \cos \alpha - \frac{\partial R}{\partial x} \cos \beta \right) d\sigma = \int_L R dz \quad (C)$$

To this end the equation of the surface should be written in the forms suitable for (B) and (C). Let the reader derive these formulas.

On adding together equalities (A), (B) and (C) we complete the derivation of Stokes' formula.

Stokes' formula can be used for reducing an integral over a surface to the corresponding line integral round the boundary of the surface and, conversely, for reducing a line integral round a closed contour to the corresponding surface integral over a surface "pulled over" the contour of integration, that is, having this contour as boundary.

In Sec. 141 Green's formula was applied to prove Theorem I on path-independence of a line integral over a plane contour. Stokes' formula can be similarly used for proving the path-independence theorem for line integrals over space curves (in Sec. 143 we did not present the proof of this theorem).

Consider the integral

$$\int P dx + Q dy + R dz \quad (*)$$

We shall suppose that the equalities

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}, \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}, \quad \frac{\partial R}{\partial x} = \frac{\partial P}{\partial z}$$

are satisfied identically in a simply connected domain Ω ; then, for an arbitrary closed contour L in Ω , we can take a surface with the contour L as boundary lying within the domain Ω and apply Stokes' theorem to prove that line integral (*) is equal to zero for any such contour L .

Conversely, let it be known that line integral (*) turns into zero for any closed contour L in Ω . Suppose that there is a point

$M_0(x_0, y_0, z_0)$ in the domain Ω for which the above equalities are violated; for definiteness, let

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \mu > 0$$

at the point M_0 . Let us draw through the point M_0 the plane $z = z_0$ parallel to the plane Oxy . Then there exists a neighbourhood of the point M_0 in that plane parallel to Oxy such that $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} > 0$ at its every point.

For the plane $z = z_0$ we have $dz = 0$, $\cos \alpha = 0$, $\cos \beta = 0$ and $\cos \gamma = 1$. Now, arguing by analogy with page 496, we conclude that there is a closed contour in Ω such that integral (*) taken over it is nonzero, which contradicts the hypothesis. The theorem has thus been proved.

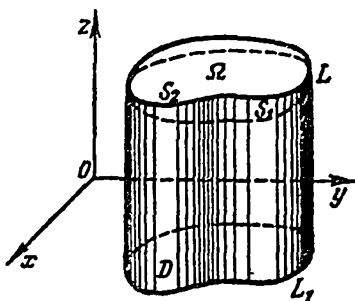


Fig. 216

150*. Gauss-Ostrogradsky¹ Formula. In this section we shall investigate a relationship between a surface integral over a closed surface and a triple integral over the domain bounded by that surface.

Gauss-Ostrogradsky Theorem. If $P(x, y, z)$, $Q(x, y, z)$ and $R(x, y, z)$ are three arbitrary functions continuous together with their partial derivatives in a bounded closed domain Ω and S is the surface bounding this domain² we have

$$\begin{aligned} \iiint_{\Omega} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \right) dx dy dz &= \\ = \iint_S [P \cos \alpha + Q \cos \beta + R \cos \gamma] d\sigma &\quad (*) \end{aligned}$$

where the integration is extended over the outer side of the surface S ($\cos \alpha$, $\cos \beta$ and $\cos \gamma$ are the direction cosines of the outer normal to the surface S).

Relation (*) is known as the *Gauss-Ostrogradsky formula*.

Proof. We shall begin with a spatial domain Ω bounded by a closed surface S such that it has at most two common points with every straight line parallel to one of the coordinate axes Ox , Oy or Oz (see Fig. 216).

¹ Gauss, K.F. (1777-1855), a great German mathematician. Ostrogradsky, M.V. (1801-1861), a noted Russian mathematician.

² The precise statement of this theorem should include conditions similar to those in the footnote on p. 491.

Let us transform the triple integral

$$I = \iiint_{\Omega} \frac{\partial R}{\partial z} dx dy dz$$

For this purpose we draw the cylindrical surface performing the orthogonal projection of the domain Ω onto the plane Oxy . It touches the surface S along a curve L dividing S into two surfaces S_1 and S_2 each of which is intersected by any straight line parallel to Oz at no more than one point. Let the domain D be the projection of the surfaces S_1 and S_2 (and of the domain Ω) on the plane Oxy , equations $z = z_2(x, y)$ and $z = z_1(x, y)$ representing the surfaces S_2 and S_1 . Performing in the triple integral the inner integration with respect to z we write it in the form

$$I = \iiint_{\Omega} \frac{\partial R}{\partial z} dx dy dz = \iint_D dx dy \int_{z_1(x, y)}^{z_2(x, y)} \frac{\partial R}{\partial z} dz$$

which leads to

$$I = \iint_D R[x, y, z_2(x, y)] dx dy - \iint_D R[x, y, z_1(x, y)] dx dy$$

The domain D being the projection of the surfaces S_2 and S_1 on the xy -plane, the double integrals on the right-hand side express, respectively, the integrals $\iint R(x, y, z) \cos \gamma d\sigma$ taken over the upper sides of these surfaces (see Formula (***) of Sec. 148). Thus,

$$I = \iint_{+S_2} R \cos \gamma d\sigma - \iint_{+S_1} R \cos \gamma d\sigma = \iint_{+S_2} R \cos \gamma d\sigma + \iint_{-S_1} R \cos \gamma d\sigma$$

Since the upper side of S_2 and the lower side of S_1 form the outer side of the whole closed surface S we have

$$I = \iiint_{\Omega} \frac{\partial R}{\partial z} dx dy dz = \iint_S R \cos \gamma d\sigma \quad (A)$$

The surface S serving as the boundary of the domain Ω may contain some cylindrical parts with elements perpendicular to the xy -plane; formula (A) applies to this case as well.

As in all similar cases in the foregoing sections, formula (A) also remains valid for domains of a more general form when these domains can be divided into subdomains satisfying the above requirements.

Arguing in a similar manner we readily derive the formulas

$$\iiint_{\Omega} \frac{\partial Q}{\partial y} dx dy dz = \iint_S Q \cos \beta d\sigma \quad (B)$$

and

$$\iiint_{\Omega} \frac{\partial P}{\partial x} dx dy dz = \iint_S P \cos \alpha d\sigma \quad (C)$$

On adding termwise equalities (A), (B) and (C) we arrive at the *Gauss-Ostrogradsky formula* (*). The theorem has been proved.

The Gauss-Ostrogradsky formula makes it possible to replace a triple integral by the corresponding surface integral over the surface bounding the domain of integration and, conversely, to replace an integral over a closed surface by the corresponding triple integral over the domain bounded by that surface.

Example. Let us apply the Gauss-Ostrogradsky formula to compute the integral $I = \iint_S (x \cos \alpha + y \cos \beta + z \cos \gamma) d\sigma$ where S is the outer side of the sphere $x^2 + y^2 + z^2 = R^2$.

By the Gauss-Ostrogradsky formula, we have

$$I = \iiint_{\Omega} (1 + 1 + 1) dx dy dz = 3V = 4\pi R^3$$

since the volume of a ball of radius R is equal to $V = \frac{4}{3} \pi R^3$.

§ 3*. Field Theory

151*. Vector Field and Vector Lines.

1. **Vector field.** The definition of a vector field is in many respects similar to the definition of a scalar field stated in Sec. 125; the reader is advised to reread § 4 of Chapter VII before starting to study this section.

Definition. If at each point P of a domain D a definite vector is specified we say that there is a *vector field* in the domain D .

We have already dealt with some concrete examples of physical vector fields. In § 1 we considered a *field of force* and studied the problem of computing the work of the field in the motion of a material point under the action of the forces of the field. In § 2 we investigated the *velocity field* of a flow of a fluid and discussed the problem of finding the fluid flux across a given surface. In § 3 we shall consider some general problems concerning vector fields. As illustrative examples of vector fields we shall most often consider force fields (such as the field of gravitation, an electric or magnetic field and the like) and velocity fields of fluid flows. These examples will demonstrate a clear physical meaning of mathematical notions introduced in the field theory.

A vector field is completely specified if the vector $\mathbf{A}(P)$ corresponding to each point P of the field is determined. We shall begin with *stationary fields* for which the vector $\mathbf{A}(P)$ is time-

independent and is only determined by the location of the point. Let A_x , A_y and A_z denote the projections of the vector $A(P)$ on the coordinate axes. If the coordinates of the point P are x , y and z the vector $A(P)$ itself and its projections are functions of these coordinates¹, and we can write

$$A(P) = A_x(x, y, z) \mathbf{i} + A_y(x, y, z) \mathbf{j} + A_z(x, y, z) \mathbf{k}$$

We shall always suppose that the functions A_x , A_y and A_z are continuous together with their partial derivatives of the first and second orders.

Let us consider some special types of vector field.

1. *Homogeneous field.* A vector field is called *homogeneous* if $A(P)$ is a constant vector, i.e. A_x , A_y and A_z are constant magnitudes.

An example of a homogeneous field is the field of gravitation studied in Sec. 144, Example 1.

2. *Plane field.* If there exists a Cartesian coordinate system $Oxyz$ in which the projections of the vectors of a field are independent of one of the variables x , y and z and one of the projections of the vectors is equal to zero the field is said to be *plane*. For instance, such is the field determined by the formula

$$A(P) = A_x(x, y) \mathbf{i} + A_y(x, y) \mathbf{j}$$

Plane fields are often dealt with in the study of *plane fluid flow*, whose feature is that all the particles of the fluid move in planes parallel to a fixed plane, the velocities of the particles at every point of any fixed straight line perpendicular to that plane being equal.

Let us discuss an important physical example. Consider a rigid body in a rotary motion with constant angular velocity ω about a fixed axis. We shall determine the field of *linear velocities* at the points of the body.

As is well known from kinematics, the linear velocity v at a point is equal to the vector product of the angular velocity ω (which is a vector lying in the axis of rotation with length equal to the absolute value ω of the angular speed and directed so that the rotation of the body, when watched from the tip of the vector, is seen counterclockwise) by the radius vector r of the variable point M of the body:

$$\mathbf{v} = \omega \times \mathbf{r}$$

In this formula the point of application of the radius vector can be taken at an arbitrary fixed point of the axis of rotation.

¹ In Sec. 70 we considered vector functions of one argument. In this section we are concerned with vector functions dependent on three scalar arguments.

Choosing this point as the origin and drawing the z -axis along the axis of rotation (in the direction of the vector ω ; see Fig. 217) we can find the projections of the vector v :

$$\omega = \omega k, \quad r = xi + yj + zk$$

which implies

$$v = \begin{vmatrix} i & j & k \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix} = -\omega yi + \omega xj$$

whence

$$v_x = -\omega y, \quad v_y = \omega x, \quad v_z = 0$$

Thus, this is a plane field.

II. Vector lines.

Definition. A *vector line* of a vector field is a curve at whose every point the direction of its tangent coincides with the direction of the vector of the field (Fig. 218).

Vector lines of concrete physical vector fields can be given a clear physical interpretation. For instance, the vector lines of a

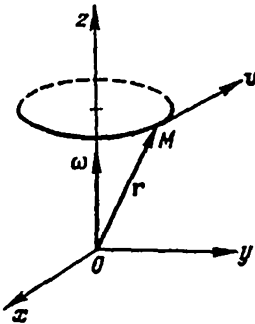


Fig. 217

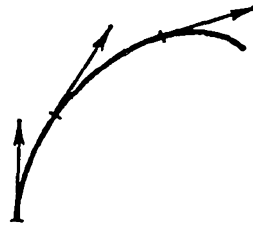


Fig. 218

stationary velocity field of a fluid flow are the so-called *stream lines* (flow lines) which serve as the trajectories of motion of the particles of the fluid. For an electric field its vector lines are the *lines of force* of the field. For example, in the electric field of a point charge these lines are the rays issued from the point at which the charge is located. The vector lines of a magnetic field of a magnet (the lines of force of that field) are curves going from the north pole to the south pole. The study of the shape and disposition of the lines of force of electric, magnetic and electromagnetic fields is extremely important for physics.

Let us derive equations of vector lines. Let a function

$$\mathbf{A}(P) = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$$

specify a vector field (for the sake of brevity, we do not indicate the arguments x, y and z the functions A_x, A_y and A_z depend on). If a vector line is represented by parametric equations

$$x = x(t), \quad y = y(t), \quad z = z(t)$$

the projections of the direction vector of the tangent line to that curve are proportional to the derivatives $x'(t)$, $y'(t)$ and $z'(t)$ or, which is the same, to the differentials dx , dy and dz (see Sec. 68). Writing down the parallelism conditions for the vector $\mathbf{A}(P)$ and the vector with components dx, dy, dz directed along the tangent to the vector line we obtain

$$\frac{dx}{A_x} = \frac{dy}{A_y} = \frac{dz}{A_z} \quad (*)$$

Relations (*) are a *system of differential equations* of the family of vector lines of the field $\mathbf{A}(P)$. Systems of differential equations will be studied later (see Sec. 176), and for the time being we can only consider vector lines of a field for some simple cases whose physical meaning allows us to judge upon the properties of the vector lines. We also note that if the field is *plane*, that is, if, for instance, $A_z = 0$ and A_x, A_y only depend on x and y , its vector lines are in the planes parallel to the plane Oxy , and their equations reduce to the form $\frac{dx}{A_x} = \frac{dy}{A_y}$.

Example. Let us determine the vector lines of the field of linear velocities of a body rotating with constant angular velocity about Oz .

According to Example 1, the function representing this field is written as

$$\mathbf{A}(P) = -\omega y \mathbf{i} + \omega x \mathbf{j}$$

Fig. 217 clearly shows that the vector lines of this field are circles lying in the planes parallel to the plane Oxy with centre on the z -axis. The equations of these circles are

$$x^2 + y^2 = R, \quad z = z_0 = \text{const}$$

In this case it is readily checked that the equality $\frac{dx}{-\omega y} = \frac{dy}{\omega x}$ holds. Indeed, the differentiation of the equation of the circle yields $2x dx + 2y dy = 0$, that is, $\frac{dx}{-y} = \frac{dy}{x}$. On multiplying the denominators of this equality by ω we arrive at the desired result.

152*. Flux of a Vector Field. Divergence.

1. Flux of a vector field. Let there be given a vector field

$$\mathbf{A}(P) = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$$

Consider a surface S in this field on which a definite side is chosen¹. Let $\mathbf{n}$ be unit normal vector to the chosen side of the surface at its variable point P ; the projections of the vector $\mathbf{n}$ are its direction cosines, that is, $\cos\alpha$, $\cos\beta$ and $\cos\gamma$. Let us form the integral of the scalar product of the vector $\mathbf{A}(P)$ by the unit vector $\mathbf{n}$ over the surface S :

$$\iint_S \mathbf{A}(P) \mathbf{n} \, d\sigma = \iint_S (A_x \cos\alpha + A_y \cos\beta + A_z \cos\gamma) \, d\sigma \quad (*)$$

If $\mathbf{A}(P)$ is the velocity field of a fluid flow integral (*) expresses the *flux of the fluid* through the surface S . If $\mathbf{A}(P)$ is an arbitrary vector field, integral (*) is called the *flux of the vector field* $\mathbf{A}(P)$ (or, briefly, the *flux of vector* $\mathbf{A}(P)$) across the surface S ; we shall denote it by the letter K .

Definition. The *flux of a vector field* across a surface is the integral of the scalar product of the vector of the field by the unit normal vector to the surface taken over that surface:

$$K = \iint_S \mathbf{A}(P) \mathbf{n} \, d\sigma$$

Thus, the computation of the flux of a vector reduces to the computation of the corresponding surface integral. The definition implies that the flux K of the vector field is a scalar magnitude. If the direction of the normal $\mathbf{n}$ is reversed, that is, if the side of the surface S is changed, the flux K changes sign. Since the scalar product of the vector $\mathbf{A}(P)$ by unit normal vector $\mathbf{n}$ is equal to the projection $A_n(P)$ of the vector $\mathbf{A}(P)$ on the direction of $\mathbf{n}$ the flux K can be written in the form

$$K = \iint_S A_n(P) \, d\sigma$$

In particular, it follows that if the projection of the vector $\mathbf{A}(P)$ on the normal is constant for a part of the surface, that is, $A_n(P) = A_n = \text{const}$, the flux through this portion of the surface is simply equal to $A_n Q$ where Q is the area of that part of the surface.

Example. Let us compute the flux of the radius vector $\mathbf{r}$ across the lateral surface S_1 of the right circular cylinder of radius R and altitude H with Oz as axis and centre of the lower base at the origin and also the flux across its upper and lower bases S_2

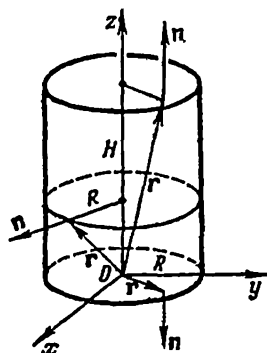


Fig. 219

¹ We only deal with two-sided surfaces S (see p. 520).

and S_3 (Fig. 219). The vector $\mathbf{n}$ is the outer normal on all the parts of the surface of the cylinder.

The normal $\mathbf{n}$ to the lateral surface S_1 is parallel to the plane Oxy , and the projection r_n is equal to R . Therefore,

$$K_1 = \iint_{S_1} R d\sigma = R \cdot 2\pi R H = 2\pi R^2 H$$

The outer normal $\mathbf{n}$ to the upper base goes along Oz , and $r_n = H$ for it. Consequently,

$$K_2 = \iint_{S_2} H d\sigma = \pi R^2 H$$

Finally, for the lower base S_3 we have $r_n = 0$ and $K_3 = 0$.

The case of an integral over a closed surface S bounding a domain Ω is of particular interest. If the integral is taken over the *outer* side, that is, if $\mathbf{n}$ is the outer normal, we speak of a *flux in the outer direction*. For this case we shall denote the flux as

$$K = \oiint_S A_n(P) d\sigma$$

Generally, in vector analysis and its applications integrals over closed surfaces and over closed contours are designated by the symbols $\oiint_S$ and $\oint_L$. It should also be noted that in physics and

applied sciences surface integrals (and double integrals as their special case for plane surfaces) and even triple integrals are often denoted by means of one sign of integration $\int$. In these symbols the type of the integral is recognized by the notation of the domain of integration: $\int_L f(M) ds$ designates a *line integral*,

$\int_S f(M) d\sigma$ designates a *surface integral* and $\int_\Omega f(M) dv$ a triple integral over a *spatial domain*.

If $A(P)$ is a velocity field of a fluid flow the magnitude of the flux K is equal to the difference between the amounts of fluid flowing out of the domain Ω and flowing into it. If $K = 0$ these amounts of fluid are in a balance for the domain Ω . For instance, the latter situation takes place for any domain isolated (mentally) from the water flow in a river.

If the magnitude K is nonzero, for instance, positive, the out-flowing amount of fluid from the domain Ω exceeds the amount of fluid flowing into the domain. This means that there are *sources* inside the domain Ω generating the fluid. Conversely, if the magnitude K is negative there are *sinks* (also termed *negative sources*) inside the domain Ω where the fluid disappears.

II. Divergence. Let us take a point P in a vector field $A(P)$ and consider a closed surface S entirely lying within the domain occupied by the field and containing the point P in its interior. Let us compute the flux of the field across the surface S and form the ratio of that flux to the volume V of the domain Ω bounded by the surface S :

$$\frac{\oiint_S A_n(P) d\sigma}{V}$$

For a fluid flow this ratio is equal to the amount of fluid generated or disappearing within the domain in unit time related to unit volume; as we say, it characterizes the *average (volume) source intensity*; if the flux across the surface S in the direction of outer normal is negative we speak of the *sink intensity*.

Now let us find the limit

$$\lim \frac{\oiint_S A_n(P) d\sigma}{V}$$

of this ratio as the domain Ω is contracted to the point P and its volume V tends to zero. If this limit exists and is positive we say that there is a *point source* of fluid at the point P ; if the limit is negative we speak of a *sink* at the point P . The magnitude of this limit characterizes the *source or sink intensity* at the point P . In the former case the fluid is generated in any infinitesimal volume containing the point P inside it and in the latter case the fluid disappears in every such volume. The limit itself is called the *divergence* of the vector field at the point P .

Definition. The *divergence* of a vector field $A(P)$ at a point P is the limit of the ratio of the flux of the vector field across a closed surface enclosing the point P to the volume of the domain bounded by the surface on condition that the surface is contracted to the point P .

The divergence is denoted by the symbol $\operatorname{div} A(P)$. Thus,

$$\operatorname{div} A(P) = \lim \frac{\oiint_S A_n(P) d\sigma}{V}$$

where the limit is computed on condition that the surface S is contracted to the point P .

We shall prove that under the requirements imposed on the field (i.e. the continuity of the functions A_x , A_y , A_z and of their first and second partial derivatives) the divergence of the field, that is, the above limit, exists for every point of the field. We

shall also compute the divergence which, as should be stressed, is a *scalar magnitude*.

Theorem. The divergence of the vector field

$$\mathbf{A}(P) = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$$

at a point P is expressed by the formula

$$\operatorname{div} \mathbf{A}(P) = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

where the values of the partial derivatives are taken at the point P .

Proof. By Gauss-Ostrogradsky formula (see Sec. 150), the flux K of the vector field through a closed surface S enclosing a given point P is representable in the form

$$\begin{aligned} \oiint_S A_n(P) d\sigma &= \oiint_S (A_x \cos \alpha + A_y \cos \beta + A_z \cos \gamma) d\sigma = \\ &= \iiint_{\Omega} \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx dy dz \end{aligned}$$

where Ω is the neighbourhood of the point P bounded by the surface S . Using the mean-value theorem (see Sec. 133) we can write the triple integral on the right as the product of the volume V by the value of the integrand at an intermediate point P_1 of the domain Ω :

$$\iiint_{\Omega} \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx dy dz = \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right)_{P_1} V$$

When the domain Ω is contracted to the point P the point P_1 tends to the point P , and we obtain

$$\operatorname{div} \mathbf{A}(P) = \lim_{P_1 \rightarrow P} \frac{\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right)_{P_1} V}{V} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

which is what we had to prove.

Using the expression obtained for the divergence we can restate the Gauss-Ostrogradsky theorem as follows:

The flux of a vector field through a closed surface in the direction of the outer normal is equal to the triple integral of the divergence of the field over the volume bounded by this surface:



$$\oiint_S A_n(P) d\sigma$$

tion expresses ti

locity field of fl
theorem ex

$$\mathbf{A}(P) d\mathbf{v}$$

ogradsky f.

vector for
obvious f.



or

SS-

across a surface is equal to the resultant intensity of all the sources and sinks enclosed by the surface, that is, to the amount of fluid flowing out of the domain bounded by the surface or flowing into it in unit time. (If the intensity of sinks exceeds that of the sources the fluid disappears within the given domain.) In particular, if the divergence is equal to zero at all the points the flux through any closed surface is equal to zero.

Examples. (1) Consider a *homogeneous* field

$$\mathbf{A}(P) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

where a , b and c are constants. It is obvious that the divergence of this field is identically equal to zero:

$$\operatorname{div} \mathbf{A}(P) \equiv 0$$

If all the particles of fluid have the same velocity, that is, if the fluid undergoes a plane-parallel displacement, there must be neither sources nor sinks in that flow. Accordingly, the fluid flux across any closed surface is equal to zero for such a motion.

(2) Let us compute the divergence of the field of linear velocities of a rotating body (see Example 1 of Sec. 151). This field is written

$$\mathbf{v} = -\omega y\mathbf{i} + \omega x\mathbf{j}$$

and its divergence is

$$\operatorname{div} \mathbf{v} = \frac{\partial(-\omega y)}{\partial x} + \frac{\partial(\omega x)}{\partial y} = 0$$

If we imagine a volume of fluid rotating as a rigid body it becomes clear that there are no sources and sinks in such a flow.

(3) Let us find the divergence of the radius vector $\mathbf{r}$. We have

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

and therefore

$$\operatorname{div} \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$$

Each point of this field serves as a source of constant intensity. Using the Gauss-Ostrogradsky formula in vector form we immediately find that the flux of the radius vector $\mathbf{r}$ through any closed surface is three times the volume bounded by the surface:

$$K = \iiint_{\Omega} 3dV = 3V$$

III. Properties of the divergence. When computing the divergence it is convenient to use its general properties enumerated below.

1. $\operatorname{div}[C_1\mathbf{A}_1(P) + C_2\mathbf{A}_2(P)] = C_1\operatorname{div} \mathbf{A}_1(P) + C_2\operatorname{div} \mathbf{A}_2(P)$ where C_1 and C_2 are scalar constants.

shall also compute the divergence which, as should be stressed, is a *scalar magnitude*.

Theorem. The divergence of the vector field

$$\mathbf{A}(P) = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$$

at a point P is expressed by the formula

$$\operatorname{div} \mathbf{A}(P) = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

where the values of the partial derivatives are taken at the point P .

Proof. By Gauss-Ostrogradsky formula (see Sec. 150), the flux K of the vector field through a closed surface S enclosing a given point P is representable in the form

$$\begin{aligned} \oiint_S A_n(P) d\sigma &= \oiint_S (A_x \cos \alpha + A_y \cos \beta + A_z \cos \gamma) d\sigma = \\ &= \iiint_{\Omega} \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx dy dz \end{aligned}$$

where Ω is the neighbourhood of the point P bounded by the surface S . Using the mean-value theorem (see Sec. 133) we can write the triple integral on the right as the product of the volume V by the value of the integrand at an intermediate point P_1 of the domain Ω :

$$\iiint_{\Omega} \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) dx dy dz = \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right)_{P_1} V$$

When the domain Ω is contracted to the point P the point P_1 tends to the point P , and we obtain

$$\operatorname{div} \mathbf{A}(P) = \lim_{P_1 \rightarrow P} \frac{\left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right)_{P_1} V}{V} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$$

which is what we had to prove.

Using the expression obtained for the divergence we can restate the Gauss-Ostrogradsky theorem as follows:

The flux of a vector field through a closed surface in the direction of the outer normal is equal to the triple integral of the divergence of the field over the spatial domain bounded by this surface:

$$\oiint_S A_n(P) d\sigma = \iiint_{\Omega} \operatorname{div} \mathbf{A}(P) d\sigma$$

This relation expresses the Gauss-Ostrogradsky formula in *vector form*.

For a velocity field of fluid flow the vector form of the Gauss-Ostrogradsky theorem expresses the obvious fact that the flux

across a surface is equal to the resultant intensity of all the sources and sinks enclosed by the surface, that is, to the amount of fluid flowing out of the domain bounded by the surface or flowing into it in unit time. (If the intensity of sinks exceeds that of the sources the fluid disappears within the given domain.) In particular, if the divergence is equal to zero at all the points the flux through any closed surface is equal to zero.

Examples. (1) Consider a *homogeneous* field

$$\mathbf{A}(P) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$$

where a , b and c are constants. It is obvious that the divergence of this field is identically equal to zero:

$$\operatorname{div} \mathbf{A}(P) \equiv 0$$

If all the particles of fluid have the same velocity, that is, if the fluid undergoes a plane-parallel displacement, there must be neither sources nor sinks in that flow. Accordingly, the fluid flux across any closed surface is equal to zero for such a motion.

(2) Let us compute the divergence of the field of linear velocities of a rotating body (see Example 1 of Sec. 151). This field is written

$$\mathbf{v} = -\omega y\mathbf{i} + \omega x\mathbf{j}$$

and its divergence is

$$\operatorname{div} \mathbf{v} = \frac{\partial(-\omega y)}{\partial x} + \frac{\partial(\omega x)}{\partial y} = 0$$

If we imagine a volume of fluid rotating as a rigid body it becomes clear that there are no sources and sinks in such a flow.

(3) Let us find the divergence of the radius vector $\mathbf{r}$. We have

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

and therefore

$$\operatorname{div} \mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$$

Each point of this field serves as a source of constant intensity. Using the Gauss-Ostrogradsky formula in vector form we immediately find that the flux of the radius vector $\mathbf{r}$ through any closed surface is three times the volume bounded by the surface:

$$K = \iiint_{\Omega} 3dV = 3V$$

III. Properties of the divergence. When computing the divergence it is convenient to use its general properties enumerated below.

1. $\operatorname{div}[C_1\mathbf{A}_1(P) + C_2\mathbf{A}_2(P)] = C_1\operatorname{div}\mathbf{A}_1(P) + C_2\operatorname{div}\mathbf{A}_2(P)$ where C_1 and C_2 are scalar constants.

For, if the projections of the vectors $A_1(P)$ and $A_2(P)$ are denoted, respectively, A_{1x}, A_{1y}, A_{1z} and A_{2x}, A_{2y}, A_{2z} , we have

$$\begin{aligned} C_1 A_1(P) &= C_1 A_1(P) = \\ &= (C_1 A_{1x} + C_1 A_{2x}) i + (C_1 A_{1y} + C_1 A_{2y}) j + (C_1 A_{1z} + C_1 A_{2z}) k \end{aligned}$$

Therefore,

$$\begin{aligned} \operatorname{div} [C_1 A_1(P) + C_2 A_2(P)] &= C_1 \frac{\partial A_{1x}}{\partial x} + C_2 \frac{\partial A_{2x}}{\partial x} + C_1 \frac{\partial A_{1y}}{\partial y} + C_2 \frac{\partial A_{2y}}{\partial y} + \\ &+ C_1 \frac{\partial A_{1z}}{\partial z} + C_2 \frac{\partial A_{2z}}{\partial z} = C_1 \operatorname{div} A_1(P) + C_2 \operatorname{div} A_2(P) \end{aligned}$$

2. Let $A(P)$ be a vector function specifying a vector field and $u(P)$ a scalar function specifying a scalar field. Then

$$\operatorname{div} [u(P) A(P)] = u(P) \operatorname{div} A(P) + A(P) \operatorname{grad} u(P)$$

Indeed, we have

$$u(P) A(P) = u A_x i + u A_y j + u A_z k$$

whence

$$\begin{aligned} \operatorname{div} [u(P) A(P)] &= \frac{\partial (u A_x)}{\partial x} + \frac{\partial (u A_y)}{\partial y} + \frac{\partial (u A_z)}{\partial z} = \\ &= u \left(\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right) + \left(A_x \frac{\partial u}{\partial x} + A_y \frac{\partial u}{\partial y} + A_z \frac{\partial u}{\partial z} \right) \end{aligned}$$

Since the gradient of u is expressed as $\frac{\partial u}{\partial x} i + \frac{\partial u}{\partial y} j + \frac{\partial u}{\partial z} k$, the expression in the second bracket is the scalar product of the vector $A(P)$ of the field by the gradient of the function $u(P)$; the formula has been proved.

In particular (see Example 3), we have

$$\operatorname{div} u r = 3u + r \operatorname{grad} u$$

153*. Circulation and Rotation of a Vector Field.

I. Circulation. Let there be a vector field described by the formula

$$A(P) = A_x i + A_y j + A_z k$$

Take a curve L in the domain occupied by the field and choose a definite direction on it. We shall denote by ds the vector going along the tangent to the curve L in the chosen direction and having the absolute value equal to the differential of arc length ds . Then (see Sec. 68) we have

$$ds = dx i + dy j + dz k$$

Let us consider the line integral of the scalar product of the

vectors $A(P)$ and ds taken along the curve L :

$$\int_L A(P) ds = \int_L A_x dx + A_y dy + A_z dz \quad (*)$$

If $A(P)$ is a field of force integral (*) expresses the work of the field performed when the material point moves in the curve L (see Sec. 143).

If $A(P)$ is an arbitrary vector field and L a closed contour, integral (*) is termed the *circulation of the vector field* $A(P)$ (or, simply, of the *vector* $A(P)$) *over the contour* L .

Definition. The *circulation of a vector field* $A(P)$ over a closed contour L is the line integral of the scalar product of the vector $A(P)$ by the vector ds tangent to the contour taken round that contour.

Since the scalar product $A(P)ds$ is equal to $A_s(P)ds$ where $A_s(P)$ is the projection of the vector $A(P)$ on the direction of the tangent line to the contour and ds is the differential of arc length the circulation can be written as the line integral with respect to arc length (see Sec. 145):

$$\oint_L A_s(P) ds$$

We remind the reader that if the element of integration in the integral (*) is a total differential and the vector field occupies a simply connected domain Ω (see p. 505) this integral turns into zero for any closed contour (see Sec. 143).

For an arbitrary vector field the circulation is a number dependent on the contour L . For instance, if there are closed vector lines in the given vector field (see the example on p. 536) and the contour L coincides with such a vector line (in its position and direction) we have $A_s(P) = |A(P)|$, and, consequently, the circulation must be a positive number as integral of a positive function. If the direction of integration is changed to the opposite the circulation becomes negative. If L is not a vector line then, the closer the directions of the vectors of the field $A(P)$ to the directions of the tangents to the vector lines, the greater is the value of the circulation.

Let us compute the circulation of the vector field of linear velocities of a rotating body (see Sec. 151)

$$\mathbf{v} = -\omega y \mathbf{i} + \omega x \mathbf{j}$$

over an arbitrary contour L lying entirely in a plane Π whose normal

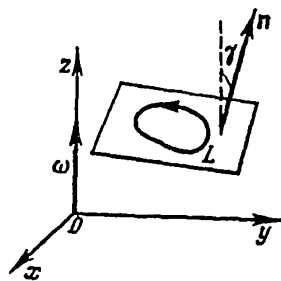


Fig. 220

vector $\mathbf{n}$ forms some angles α , β and γ with the coordinate axes (see Fig. 220). The direction in which the contour L is traversed is coherent with the direction of the normal $\mathbf{n}$ (as in Stokes' theorem; see Sec. 149). Using the definition of circulation we find that for such a contour it is equal to

$$\oint_L -\omega y dx + \omega x dy$$

To compute this integral we apply Stokes' theorem:

$$\omega \oint_L -y dx + x dy = \omega \iint_D 2 dx dy$$

where D is the domain bounded by the contour L in the plane Π . The integration goes over the upper side of the plane Π , and therefore we have $dx dy = \cos \gamma d\sigma$ (see Sec. 148). Thus, since $\cos \gamma = \text{const}$, we obtain the expression

$$2\omega \cos \gamma \iint_D d\sigma = 2\omega \cos \gamma S$$

for the circulation where S is the area of the domain bounded by the contour L . Now, note that $\omega \cos \gamma$ is equal to the projection ω_n of the vector ω on the direction of the vector $\mathbf{n}$. Therefore we arrive at the following expression for the circulation:

$$\oint_L -\omega y dx + \omega x dy = 2\omega_n S$$

(If L is a circle of radius R the circulation is equal to $2\omega_n \pi R^2$.)

The latter formula implies that if the plane Π is rotated, that is, if the angle γ is changed, the value of the circulation becomes variable. It attains its greatest value for $\gamma=0$, that is, for the plane Π parallel to the plane Oxy with normal vector $\mathbf{n}$ parallel to the angular velocity ω . If $\gamma = \frac{\pi}{2}$, i.e. if the normal to the plane is perpendicular to the vector ω , the circulation turns into zero.

It is also clear that if we choose the other side of the plane Π the vector $\mathbf{n}$ changes its direction to the opposite and the projection ω_n becomes negative. Hence, in this case the circulation is also negative.

Let us discuss the physical meaning of the circulation of a vector in the case when $\mathbf{A}(P)$ is a velocity field of fluid flow. For the sake of simplicity, we shall assume that the contour L is a circle lying in a plane. Suppose that this circle is the periphery of a turbine with plane vanes parallel to the axis of the turbine (this fixed axis about which the turbine can rotate is perpendicular to the plane of the turbine and passes through its centre).

If the circulation is equal to zero the turbine remains motionless since the resultant of the forces acting upon it is equal to zero. If the circulation is nonzero the turbine rotates, and, the greater its value, the higher the angular velocity. For instance, if the fluid rotates as a rigid body about the z -axis and the axis of rotation coincides with the axis of the turbine then, as has been shown, the circulation is equal to $2\omega S$ where S is the area of the turbine. Hence, the ratio of the circulation to the area of the turbine is twice the angular speed and is independent of the dimensions of the turbine.

If the axis of the turbine forms an angle with Oz this ratio decreases: it becomes equal to $2\omega_n$ where ω_n is the projection of the vector ω on the axis of the turbine. Finally, if the axis of the turbine is perpendicular to the axis of rotation of the fluid the turbine is obviously in the state of rest.

For an arbitrary vector field the ratio of the circulation over a plane contour L to the area S bounded by the contour is a *variable quantity*. To characterize rotational properties of a vector field at its point P it is natural to form the ratio of the circulation over a plane contour L enclosing the point P to the area S bounded by the contour and consider the limit of this ratio as the contour L contracts to the point P remaining in the fixed plane:

$$\lim_{\substack{L \\ S}} \frac{\oint A_s ds}{S}$$

We shall show that this limit depends solely on the choice of the point P and on the direction of the normal to the plane in which the contour L lies. In this construction the direction of the normal must be coherent with the direction in which the contour is described: if the contour is watched from the terminus of the normal it is described counterclockwise. In order to compute this limit we transform the expression of the circulation with the aid of Stokes' formula (see Sec. 149):

$$\begin{aligned} \oint_L A_s ds &= \oint_L A_x dx + A_y dy + A_z dz = \\ &= \iint_D \left[\left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \cos \alpha + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \cos \beta + \right. \\ &\quad \left. + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \cos \gamma \right] d\sigma \end{aligned}$$

where $\cos \alpha$, $\cos \beta$ and $\cos \gamma$ are the direction cosines of the normal $\mathbf{n}$ and D is the plane region bounded by the contour L . By the mean value theorem, the last integral is equal to the product

of the value of the integrand at a point P_1 of the domain D by the area S of this domain.

When the contour L is contracted to the point P the integrand tends to its value assumed at the point P . Therefore we obtain

$$\lim_L \frac{\oint A_s ds}{S} = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \cos \alpha + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \cos \beta + \\ + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \cos \gamma$$

where the values of the partial derivatives are taken at the point P . The right-hand side of this relation is a scalar product of two vectors: of the unit normal vector $\mathbf{n}$ to the plane of the contour L (its projections on the coordinate axes are $\cos \alpha$, $\cos \beta$ and $\cos \gamma$) and the vector with projections

$$\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z}, \quad \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \quad \text{and} \quad \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y}$$

The latter vector is called the *rotation* (or *curl*) of the vector field $\mathbf{A}(P)$ (or, simply, of the vector $\mathbf{A}(P)$) and is denoted $\text{rot } \mathbf{A}(P)$ (or $\text{curl } \mathbf{A}(P)$).

II. Rotation of a vector field and its properties.

Definition. The *rotation* of a vector field.

$$\mathbf{A}(P) = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$$

is the vector (vector field) specified by the formula

$$\text{rot } \mathbf{A}(P) = \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \mathbf{k}$$

The projection $\text{rot}_n \mathbf{A}(P)$ of the vector $\text{rot } \mathbf{A}(P)$ on an arbitrary direction $\mathbf{n}$ is equal to the limit of the ratio of the circulation of the vector field $\mathbf{A}(P)$ over a contour enclosing the point P and lying in the plane passing through the point P and having the vector $\mathbf{n}$ as normal to the area bounded by the contour, on condition that the contour is contracted to the point P in that fixed plane. This limit attains its greatest value when the directions of the normal $\mathbf{n}$ and of the vector $\text{rot } \mathbf{A}(P)$ coincide.

Using the definition of the rotation we can restate Stokes' theorem as follows:

The flux of the rotation of a vector field through a surface S is equal to the circulation of the vector field over the boundary of the surface:

$$\iint_S \text{rot}_n \mathbf{A}(P) d\sigma = \oint_L A_s(P) ds$$

The latter relation expresses Stokes' formula in *vector form*.

It follows that the fluxes of the rotation across any two surfaces of the type of S having a given contour L as boundary are always equal.

Let the reader verify the following property of the rotation of a vector field:

$$\text{rot}[C_1 \mathbf{A}_1(P) + C_2 \mathbf{A}_2(P)] = C_1 \text{rot} \mathbf{A}_1(P) + C_2 \text{rot} \mathbf{A}_2(P)$$

where C_1 and C_2 are scalar constants.

We shall also prove that if $u(P)$ is a scalar function and $\mathbf{A}(P)$ a vector function then

$$\text{rot } u\mathbf{A} = u \text{rot } \mathbf{A} + (\text{grad } u \times \mathbf{A})$$

To this end we shall show that the projections of the vectors on the left-hand and right-hand sides of this relation on the coordinate axes coincide. For definiteness, let us take the axis Ox . Since the projections of the vector $u\mathbf{A}$ on the axes Ox , Oy and Oz are uA_x , uA_y and uA_z , we have

$$\text{rot}_x u\mathbf{A} = \frac{\partial u A_z}{\partial y} - \frac{\partial u A_y}{\partial z} = u \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \left(A_z \frac{\partial u}{\partial y} - A_y \frac{\partial u}{\partial z} \right).$$

The first term on the right is nothing but $u \text{rot}_x \mathbf{A}$. Since

$$\text{grad } u \times \mathbf{A} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ A_x & A_y & A_z \end{vmatrix}$$

we have

$$(\text{grad } u \times \mathbf{A})_x = A_z \frac{\partial u}{\partial y} - A_y \frac{\partial u}{\partial z}$$

the latter expression being the second term on the right-hand side of the desired formula. The equality of the projections on the other coordinate axes is proved similarly.

154*. Hamiltonian¹ Operator. Repeated Application of the Hamiltonian Operator to a Vector Field. The fundamental notions of *gradient*, *divergence* and *rotation* introduced in vector analysis can be represented in a convenient form by means of a special symbolic vector operator ∇ (introduced by W.R. Hamilton) called *nabla* or *del* or the *Hamiltonian operator*. It is defined by the formula

$$\nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k}$$

¹ Hamilton, W.R. (1805-1865), an Irish mathematician.

where $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ symbolize the differentiation with respect to x , y and z . The multiplication of $\frac{\partial}{\partial x}$ by an expression (scalar or vector) is understood as the differentiation of this expression with respect to x ; the symbols $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ are used similarly.

Let us enumerate the operation rules for this vector operator:

1. The product of the vector nabla by a scalar function $u(P)$ results in the gradient of this function:

$$\nabla u = \left(\frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \right) u = \frac{\partial u}{\partial x} \mathbf{i} + \frac{\partial u}{\partial y} \mathbf{j} + \frac{\partial u}{\partial z} \mathbf{k} = \text{grad } u$$

2. The scalar product of the vector ∇ by a vector function $\mathbf{A}(P)$ results in the divergence of this vector function:

$$\begin{aligned} \nabla \cdot \mathbf{A}(P) &= \left(\frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \right) (A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}) = \\ &= \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = \text{div } \mathbf{A}(P) \end{aligned}$$

3. The vector product of the vector ∇ by a vector function $\mathbf{A}(P)$ is equal to the rotation of the vector field $\mathbf{A}(P)$:

$$\begin{aligned} \nabla \times \mathbf{A}(P) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A_x & A_y & A_z \end{vmatrix} = \\ &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \mathbf{i} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \mathbf{j} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \mathbf{k} = \text{rot } \mathbf{A}(P) \end{aligned}$$

Thus, operations involving the vector nabla are performed according to the ordinary rules of vector algebra under the convention that the multiplication of $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$ by an expression is understood as the differentiation of this expression with respect to the corresponding variable.

The operations of computing the gradient, the divergence and the rotation of a given field only contain partial derivatives of the first order of the scalar functions involved, i.e. the given scalar fields and the projections of the given vector fields.

The operator ∇ can also be applied repeatedly to a given field. Let us discuss the repeated operations involving ∇ .

Let $u(P)$ be a scalar field. The application of the operator ∇ to this field results in its gradient $\text{grad } u$. The vector $\text{grad } u$ forms a vector field, and we can compute its divergence and rotation ($\text{div grad } u$ and $\text{rot grad } u$).

If we have a *vector field* $A(P) = A_x \mathbf{i} + A_y \mathbf{j} + A_z \mathbf{k}$ the application of the operator ∇ generates two new fields: the *scalar field* $\text{div } A(P)$ and the *vector field* $\text{rot } A(P)$. Therefore, the operator ∇ can be applied repeatedly to obtain the gradient of the former field, i.e. $\text{grad div } A(P)$ and the divergence and the rotation of the latter field, i.e. $\text{div rot } A(P)$ and $\text{rot rot } A(P)$.

Thus, we have *five* operations involving repeated application of the vector operator ∇ . Among them the most important role is played by the following three operations which we shall discuss in more detail: $\text{div grad } u$, $\text{rot grad } u$ and $\text{div rot } A(P)$.

(a) For the operation $\text{div grad } u$ we have

$$\text{div grad } u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

Indeed, the gradient of u is given by the formula $\text{grad } u = \frac{\partial u}{\partial x} \mathbf{i} + \frac{\partial u}{\partial y} \mathbf{j} + \frac{\partial u}{\partial z} \mathbf{k}$. On forming the divergence of the vector $\text{grad } u$ we arrive at the desired equality. Its right-hand side involves the operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ called the *Laplacian¹ operator* and denoted Δ ; its application to a function u results in the expression

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}$$

Using the symbolic vector ∇ we can rewrite the expression of $\text{div grad } u = \Delta u$ as

$$\text{div grad } u = \Delta u = \nabla \cdot (\nabla u) = \nabla^2 u$$

This notation ($\nabla^2 = \Delta$) of the Laplacian operator is also commonly used.

(b) For any scalar field u we have the identity

$$\text{rot grad } u = 0$$

The identity is readily checked: in this case every component of the rotation is the difference of two mixed partial derivatives of the second order of the function u only differing in the order of differentiation and therefore is equal to zero; for instance,

$$\text{rot}_x \text{grad } u = \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial z} \right) - \frac{\partial}{\partial z} \left(\frac{\partial u}{\partial y} \right) = 0$$

To memorize this property it is convenient to write it down in the form

$$\text{rot grad } u = \nabla \times (\nabla u) = (\nabla \times \nabla) u = 0$$

using the vector ∇ . Then it becomes (formally) coherent with the rule that the vector product of two equal vectors is always zero.

¹ Laplace, P.S. (1749-1827), a French mathematician and physicist.

(c) For any vector field $A(P)$ we have the identity

$$\operatorname{div} \operatorname{rot} A(P) \equiv 0$$

Indeed, the divergence of the rotation of $A(P)$ is computed as

$$\begin{aligned} \operatorname{div} \operatorname{rot} A(P) &= \frac{\partial}{\partial x} \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) + \frac{\partial}{\partial y} \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) + \\ &+ \frac{\partial}{\partial z} \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) = \frac{\partial^2 A_z}{\partial y \partial x} - \frac{\partial^2 A_y}{\partial z \partial x} + \frac{\partial^2 A_x}{\partial z \partial y} - \frac{\partial^2 A_z}{\partial x \partial y} + \\ &+ \frac{\partial^2 A_y}{\partial x \partial z} - \frac{\partial^2 A_x}{\partial y \partial z} \end{aligned}$$

the latter expression being identically equal to zero since the mixed partial derivatives involved mutually cancel out.

Using the operator ∇ we can write this relation as

$$\operatorname{div} \operatorname{rot} A(P) = \nabla \cdot (\nabla \times A) = 0$$

In the latter form this identity is (formally) coherent with the rule that the triple scalar product of three vectors among which two vectors coincide is always equal to zero.

The other two operations $\operatorname{grad} \operatorname{div} A(P)$ and $\operatorname{rot} \operatorname{rot} A(P)$ involving repeated application of ∇ are used less frequently. We shall not write down their expressions in terms of the projections of the vector $A(P)$ and confine ourselves to indicating the following relationship between them:

$$\operatorname{rot} \operatorname{rot} A(P) = \operatorname{grad} \operatorname{div} A(P) - \Delta A(P)$$

where $\Delta A(P) = \Delta A_x i + \Delta A_y j + \Delta A_z k$ (Δ in the expression $\Delta A(P)$ symbolizes the Laplacian operator applied to the vector field $A(P)$).

155*. Properties of Some Vector Fields. In this section we shall discuss some properties of the simplest vector fields for which $\operatorname{div} A(P) = 0$ or $\operatorname{rot} A(P) = 0$ or both the divergence and the rotation are equal to zero.

I. Potential field. *A vector field*

$$A(P) = A_x i + A_y j + A_z k$$

is said to be potential if there is a function $u(x, y, z)$ such that

$$A_x = \frac{\partial u}{\partial x}, \quad A_y = \frac{\partial u}{\partial y}, \quad A_z = \frac{\partial u}{\partial z} \quad (*)$$

or, which is the same, $A = \operatorname{grad} u$.

The function $u(x, y, z)$ is called the potential or the potential function of the vector field u (see Sec. 144).

By Formula (b) of Sec. 154, we have

$$\operatorname{rot} A = \operatorname{rot} \operatorname{grad} u = 0$$

for such a field. Thus, the rotation of a potential field is identically equal to zero. That is why we call the velocity field of a

fluid flow *irrotational* when this vector field is potential; in such a flow there are no vortices in the fluid. In a potential field the circulation over any closed contour is equal to zero.

The work of a potential field (see Sec. 144) is equal to the potential difference between the initial and terminal points of the trajectory of motion of the particle:

$$\int_{(A)}^{(B)} A_x dx + A_y dy + A_z dz = u(B) - u(A)$$

The investigation of a potential field is simpler compared with the case of a field of general structure since the former is completely determined by *one* scalar function (i.e. by its potential $u(x, y, z)$). The projections of a potential vector field $A(P)$ on the coordinate axes are equal to the partial derivatives of the potential function with respect to the corresponding coordinates. In the general case a vector field $A(P)$ is determined by *three* scalar functions (the projections of the vector $A(P)$ on the coordinate axes).

If a vector field occupies a simply connected domain the results of Sec. 143 make it possible to find out whether it is potential or not. Namely:

If a vector field $A(P)$ occupies a simply connected domain and the rotation of the field is equal to zero ($\text{rot } A(P) = 0$) at all the points of the domain the given field is potential.

Indeed, if $\text{rot } A(P) = 0$ then there hold the equalities

$$\frac{\partial A_z}{\partial y} = \frac{\partial A_y}{\partial z}, \quad \frac{\partial A_x}{\partial z} = \frac{\partial A_z}{\partial x} \quad \text{and} \quad \frac{\partial A_y}{\partial x} = \frac{\partial A_x}{\partial y}$$

Consequently, there exists a function $u(x, y, z)$ whose total differential is

$$du = A_x dx + A_y dy + A_z dz$$

which is equivalent to $\text{grad } u = A(P)$; hence, $u(x, y, z)$ is a potential function of the field.

It should be noted that if a vector field is defined everywhere with the exception of some separate points its domain of definition is simply connected; hence, if its rotation is equal to zero the field is potential.

An example of a nonpotential field occupying a multiply connected domain and having the rotation identically equal to zero will be considered in Sec. 156, II.

II. Solenoidal field. A vector field $A(P)$ whose divergence is equal to zero at all points, that is,

$$\text{div } A(P) = 0,$$

is called *solenoidal*¹.

¹ From the Greek word σωλην tube. The origination of this term will be elucidated later

To characterize the properties of such a field let us consider a spatial domain Ω satisfying the following condition:

If a simple closed surface belongs to the domain Ω the subdomain bounded by this surface also belongs to Ω .¹

Examples of domains possessing this property are a ball, the interior of an ellipsoid, a cube, a half-space, etc.; a torus also is a domain belonging to this type. But it is obvious that a spherical shell does not satisfy this condition. Generally, if a domain has a cavity (even degenerated into a point) it does not possess the above property. For instance, if a metal ingot has cavities the domain occupied by it does not satisfy this condition.

Now, for a solenoidal field, we can state the following:

If a solenoidal field occupies a domain possessing the indicated property the flux of the field across any closed surface lying in this domain is equal to zero:

$$\oiint_S A_n(P) d\sigma = 0$$

This assertion is an immediate consequence of the Gauss-Ostrogradsky formula:

$$\oiint_S A_n(P) d\sigma = \iiint_V \operatorname{div} A(P) dv = 0$$

If a solenoidal field occupies a domain with cavities (which most often are separate singular points of the field) the flux of the field through a closed surface enclosing a cavity may be different from zero. An example of this kind is discussed in Sec. 156, I.

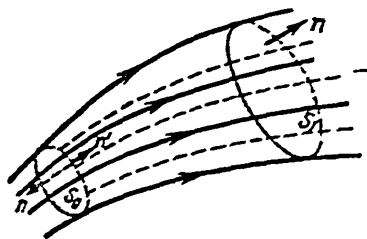


Fig. 221

Now we shall explain why a field with $\operatorname{div} A(P) = 0$ is called solenoidal, that is, why its name is connected with the notion of a tube. Take an area S_0 in the given solenoidal field and draw all vector lines of the field passing through the points of the

boundary of S_0 . These form a tubular surface enclosing a part of space called a *tube of the vector field* (Fig. 221). If this field describes a fluid flow the particles of the fluid move without intersecting the boundary of such a tube. Let us consider the part of the tube lying between the surface S_0 and a section S_1 of the

¹ This property is referred to as *acyclicity* (in dimension 2) of the domain Ω . It expresses a certain kind of "connectedness" of the domain different from that of a simply connected domain (Sec. 143). (The property of being simply connected is also referred to as *acyclicity* in dimension 1.)

tube (Fig. 221). The flux of the vector field through the whole surface bounding this part of the tube is equal to zero. Consequently,

$$\iint_S A_n d\sigma + \iint_{S_0} A_n d\sigma + \iint_{S_1} A_n d\sigma = \iiint_V \operatorname{div} \mathbf{A}(P) dv = 0$$

where S is the lateral surface of the tube and $\mathbf{n}$ is the outer normal. Since on the lateral surface the normal $\mathbf{n}$ is perpendicular to the vectors of the field, we have

$$\iint_S A_n d\sigma = 0$$

It follows that

$$\iint_{S_1} A_n d\sigma = - \iint_{S_0} A_n d\sigma$$

Let us change the direction of the normal to the surface S_0 , that is, take the inner normal $\mathbf{n}'$; this results in

$$\iint_{S_1} A_n d\sigma = \iint_{S_0} A_{n'} d\sigma$$

Hence, the flux of a solenoidal vector field in the direction of the vector lines of the field across every section of a vector tube has a constant value, that is, in a field without sources and sinks the amount of fluid flowing through every section of a vector tube is one and the same.

According to Formula (c) of Sec. 154, we have

$$\operatorname{div} \operatorname{rot} \mathbf{A}(P) = 0$$

Thus, the field of the rotation of any vector field is solenoidal.

The converse proposition which we state without proof is also true:

Every solenoidal field is a field of rotation of a vector field, i.e. if $\operatorname{div} \mathbf{A}(P) = 0$ then there exists a vector field $\mathbf{B}(P)$ such that

$$\mathbf{A}(P) = \operatorname{rot} \mathbf{B}(P)$$

The vector function $\mathbf{B}(P)$ is called a *vector potential* of the given field $\mathbf{A}(P)$.

III. Harmonic field. A vector field which is simultaneously *potential* and *solenoidal* is called *harmonic*. If $\mathbf{A}(P)$ is such a field we have

$$\mathbf{A}(P) = \operatorname{grad} u$$

where u is a potential function of the field (which exists since the field in question is supposed to be potential). The condition that the field is solenoidal means that

$$\operatorname{div} \mathbf{A}(P) = \operatorname{div} \operatorname{grad} u = 0$$

Hence, by formula (a) of Sec. 154, we have

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

A function u satisfying the latter relation (known as *Laplace's equation*) is called *harmonic*. Harmonic functions play an important role in various divisions of theoretical physics. Here we shall not dwell on them in more detail.

156*. Electromagnetic Field. There are many important applications of the notions of vector analysis which were introduced in the foregoing sections to the theory of electromagnetic fields. Here we shall discuss some simple examples of this kind.

1. Electrostatic field. Let E be the field of intensity of the electrostatic field generated by a point charge q placed at the origin. According to Example 2 of Sec. 144, the intensity of this field at a point $P(x, y, z)$ is given by the formula

$$E = \frac{qx}{r^3} \mathbf{i} + \frac{qy}{r^3} \mathbf{j} + \frac{qz}{r^3} \mathbf{k}$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the distance from the point P to the origin. The vector lines of this field are the rays issued from the origin where the charge q is placed.

The field of intensity E is potential (see Sec. 144, Example 2). In electrostatics, as the potential φ of the field E the function $u = -\frac{q}{r}$ found in Sec. 144 is usually taken with *opposite* sign:

$\varphi = \frac{q}{r}$. Therefore, $E = \text{grad} \left(-\frac{q}{r} \right) = -\text{grad} \varphi$. Hence, the potential difference between two points P_1 and P_2 of the field taken with opposite sign is equal to the work performed by the forces of the field when unit positive charge is moved from the former point to the latter.

Let us find the *divergence* of this field. The computations yield

$$\begin{aligned} \frac{\partial \left(\frac{qx}{r^3} \right)}{\partial x} &= q \frac{r - 3x \frac{\partial r}{\partial x}}{r^4} = q \frac{r^2 - 3x^2}{r^5}, & \frac{\partial \left(\frac{qy}{r^3} \right)}{\partial y} &= q \frac{r^2 - 3y^2}{r^5} \\ \frac{\partial \left(\frac{qz}{r^3} \right)}{\partial z} &= q \frac{r^2 - 3z^2}{r^5} \end{aligned}$$

whence

$$\text{div } E = \frac{q}{r^5} [3r^2 - 3(x^2 + y^2 + z^2)] = 0$$

Consequently,

the flux of the vector E across any closed surface not enclosing the origin is equal to zero.

If the origin, that is, the charge, is inside the surface this conclusion fails since the field is not defined at the origin.

Let us determine the flux of the vector $\mathbf{E}$ through the sphere of radius R with centre at the origin. On this sphere the direction of the vector $\mathbf{E}$ coincides with that of the outer normal, i.e. with the direction of the radius vector. Therefore,

$$E_n = |\mathbf{E}| = \frac{q}{R^2}$$

whence we find the flux:

$$K = \frac{q}{R^2} 4\pi R^2 = 4\pi q$$

We see that the magnitude of the flux is independent of the radius of the sphere. It can readily be shown that the flux remains the same for any closed surface enclosing the origin. To that end we take an arbitrary closed surface S of this kind¹ and consider a sphere with centre at the origin lying inside this surface (Fig. 222). Let us break up the domain bounded by the surface S into a number of

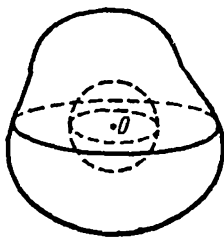


Fig. 222

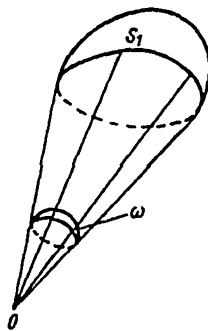


Fig. 223

cones with vertex at the origin (such a cone is shown in Fig. 223). Since the portion of each cone enclosed between the corresponding parts of the sphere and of the surface S is a tube of the vector field $\mathbf{E}$ (see Sec. 155, I), and the divergence of the field is equal to zero, the fluxes through these parts of the sphere and of the surface S are equal. On adding together the fluxes through all such parts of the sphere and of the surface S we conclude that the flux of the vector $\mathbf{E}$ across *any* closed surface containing the origin is equal to the flux across the sphere, that is, it equals $4\pi q$.

¹ The surface depicted in Fig. 222 is *star-shaped with respect to the origin*, that is, any radius vector intersects it at only one point. It can be shown that the argument of this proof also applies to the general case of an arbitrary surface.

Now suppose that the sphere taken inside the given surface has unit radius. Then the flux through the part S_1 of the surface S (see Fig. 223) is equal to $q\omega$ where ω is the area of the part of the sphere of unit radius corresponding to that part of the given surface. The quantity ω is the magnitude of the *solid angle* with vertex at the origin subtended by the surface S_1 .

Let us consider the electric field generated by a *system* of point charges $q_1, q_2, \dots, q_m$. Denoting by E_i the intensity of the field created by the charge q_i and by E the intensity of the resultant field we can write

$$E = \sum_{i=1}^m E_i$$

The projection of the vector E on the normal n to any surface is

$$E_n = \sum_{i=1}^m E_{in}$$

Thus, the flux of the field across any surface S is written as

$$K = \oint_S E_n dS = \sum_{i=1}^m \oint_S E_{in} dS = 4\pi \sum q_i$$

the last sum $\sum q_i$ being extended only over those charges which are inside the surface under consideration.

The latter formula plays an important role in the theory of electrostatic fields; in electrostatics it is known as *Gauss' theorem*¹.

Now we proceed to a *continuous* distribution of an electric charge. Let ρ denote the density of the distribution of the charge². If the charge is not distributed uniformly the density ρ is a function of the variable point P of the field, that is, a function of its coordinates. The resultant charge in a given volume Ω is equal to

$$\iiint_{\Omega} \rho dv$$

Applying Gauss' theorem to this charge we obtain

$$\iint_S E_n dS = 4\pi \iiint_{\Omega} \rho dv$$

where S is the boundary of the domain Ω .

¹ In the system of units SI the intensity of the field of a point charge q in vacuum is given by the formula $E = \frac{1}{4\pi\epsilon_0} \frac{q}{r^2} \frac{r}{r}$ where ϵ_0 is the electric constant. Accordingly, Gauss' theorem is written $K = \frac{1}{\epsilon_0} \sum q_i$.

² The definition of the density of electric charge distribution is analogous to the definition of the density of mass distribution (see Sec. 133).

Let us transform the first integral by introducing $\operatorname{div} \mathbf{E}$:

$$\iiint_{\Omega} \operatorname{div} \mathbf{E} \, dv = 4\pi \iiint_{\Omega} \rho \, dv$$

It follows that

$$\iiint_{\Omega} (\operatorname{div} \mathbf{E} - 4\pi\rho) \, dv = 0$$

Since the latter integral is equal to zero for *any* domain of integration Ω we conclude, arguing by analogy with p. 496, that

$$\operatorname{div} \mathbf{E} = 4\pi\rho \quad (*)$$

Let the potential of the electrostatic field generated by this charge distribution be φ ; then $\mathbf{E} = -\operatorname{grad} \varphi$, and, according to Formula (a) of Sec. 154, we obtain

$$\operatorname{div} \mathbf{E} = -\operatorname{div} \operatorname{grad} \varphi = -\Delta\varphi \quad (**)$$

Equalities (*) and (**) imply

$$\Delta\varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = -4\pi\rho$$

The relation obtained is called *Poisson's equation*. If the density ρ is equal to zero within a subdomain or in the whole domain in question Poisson's equation goes into Laplace's equation $\Delta\varphi = 0$ (see the end of Sec. 155). In our course we shall not study these equations.

II. Magnetic field generated by a rectilinear electric current. Let us consider the magnetic field set up by a direct electric current flowing through an infinite rectilinear conductor. We shall determine the intensity vector of the magnetic field created by this current. According to the *Biot-Savart-Laplace*¹ law, an element of current generates at a given point the magnetic field with the absolute value of intensity vector $k \frac{I \sin \alpha \, ds}{r^2}$ where I is the intensity of current, ds is the linear element of conductor, r is the distance from the element of current to the point of observation, α is the angle between the direction of the current and the straight line joining the point of observation to the element of current and k is a proportionality factor dependent on the choice of the system of units. The intensity vector goes along the normal to the plane in which the element of current and the point of observation lie, its direction being specified by *Ampère's*² rule

¹ Biot, J.B. (1775-1862) and Savart, F. (1791-1841), French physicists.

² Ampère, A.M. (1775-1836), a French physicist and mathematician.

(the *right-hand screw rule*). The vector form of the Biot-Savart-Laplace law is

$$d\mathbf{H} = k \frac{I}{r^3} (d\mathbf{s} \times \mathbf{r})$$

where $d\mathbf{H}$ is the intensity of the magnetic field created by the element of current, $d\mathbf{s}$ is the vector of length ds directed along the conductor and $\mathbf{r}$ is the vector going from the element of current to the point M at which the intensity of the field is observed.

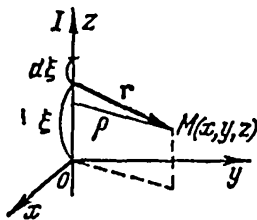


Fig. 224

Let the conductor coincide with the z -axis, the current flowing upward (see Fig. 224), and let ξ denote the variable distance from the element of current to the origin, x, y and z being the coordinates of the point M . Then

$$d\mathbf{s} = d\xi \mathbf{k}, \quad \mathbf{r} = x\mathbf{i} + y\mathbf{j} + (z - \xi)\mathbf{k}$$

and

$$r = |\mathbf{r}| = \sqrt{x^2 + y^2 + (z - \xi)^2} = \sqrt{\rho^2 + (z - \xi)^2}$$

where $\rho\sqrt{x^2 + y^2}$ is the distance from the point M to the conductor. Computing the vector product we find $d\mathbf{H}$:

$$d\mathbf{H} = \frac{kI}{r^3} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 0 & d\xi \\ x & y & z - \xi \end{vmatrix} = \frac{kI}{r^3} (-y d\xi \mathbf{i} + x d\xi \mathbf{j})$$

This means that

$$dH_x = -\frac{kI y d\xi}{(\sqrt{\rho^2 + (z - \xi)^2})^3}, \quad dH_y = \frac{kI x d\xi}{(\sqrt{\rho^2 + (z - \xi)^2})^3} \quad \text{and} \quad dH_z = 0$$

To determine H_x and H_y we should integrate the corresponding differentials from $-\infty$ to ∞ . To this end let us evaluate the improper integral

$$\int_{-\infty}^{\infty} \frac{d\xi}{(\sqrt{\rho^2 + (z - \xi)^2})^3}$$

The substitution $\xi - z = \rho \tan t$, $d\xi = \frac{\rho}{\cos^2 t} dt$ leads to the integral

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\rho dt}{\cos^2 t (\sqrt{\rho^2 (1 + \tan^2 t)})^3} = \frac{1}{\rho^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t dt = \frac{2}{\rho^2}$$

Thus,

$$H_x = -\frac{2kI y}{\rho^2} = -\frac{2kI y}{x^2 + y^2}, \quad H_y = \frac{2kI x}{\rho^2} = \frac{2kI x}{x^2 + y^2} \quad \text{and} \quad H_z = 0$$

This field is not defined at the points of the axis Oz . We see that the intensity vector $\mathbf{H}$ is in the same direction as the vector of linear velocity of a body rotating about Oz (see Sec. 151, 1) if the direction of the electric current flow coincides with that of the angular velocity. The modulus of the vector $\mathbf{H}$ is

$$|\mathbf{H}| = \sqrt{H_x^2 + H_y^2} = \frac{2kl}{\rho}$$

It is readily checked that the divergence of the field $\mathbf{H}$ is equal to zero. Indeed, we have

$$\frac{\partial H_x}{\partial x} = \frac{4lxy}{(x^2 + y^2)^2}, \quad \frac{\partial H_y}{\partial y} = -\frac{4lxy}{(x^2 + y^2)^2} \quad \text{and} \quad \frac{\partial H_z}{\partial z} = 0$$

whence

$$\operatorname{div} \mathbf{H} = 0$$

The rotation of the field is also equal to zero at all points. To verify this property it is sufficient to show that

$$\frac{\partial H_x}{\partial y} - \frac{\partial H_y}{\partial x} = 0$$

the latter being readily seen:

$$\begin{aligned} \frac{\partial H_x}{\partial y} &= 2kl \frac{y^2 - x^2}{(x^2 + y^2)^2} \quad \text{and} \\ \frac{\partial H_y}{\partial x} &= 2kl \frac{y^2 - x^2}{(x^2 + y^2)^2} \end{aligned}$$

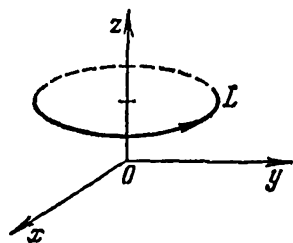


Fig. 225

Consequently, the circulation of the field over any contour *not encircling* Oz is equal to zero. But if we take a contour encircling the z -axis (see Fig. 225) this conclusion fails since any surface with such a contour as boundary intersects the axis Oz at whose points the field is not defined.

Let us compute the circulation round the circle of radius R with centre at the origin lying in the plane Oxy . This contour is represented by the equations

$$x = R \cos t, \quad y = R \sin t, \quad z = 0$$

Therefore we obtain

$$2kl \int_L \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = 2kl \int_0^{2\pi} \frac{R^2 \sin^2 t + R^2 \cos^2 t}{R^2} dt = 4\pi kl$$

The circulation is independent of the radius R of the circle. It can be proved that it retains the same value for any closed contour encircling the z -axis.

157*. Nonstationary Fields. In this section we shall discuss some properties of *nonstationary* fields for which the specifying functions

(the right-hand screw rule). The vector form of the Biot-Savart Laplace law is

$$d\mathbf{H} = k \frac{I}{r^3} (d\mathbf{s} \times \mathbf{r})$$

where $d\mathbf{H}$ is the intensity of the magnetic field created by the element of current, $d\mathbf{s}$ is the vector of length ds directed along the conductor and $\mathbf{r}$ is the vector going from the element of current to the point M at which the intensity of the field is observed.

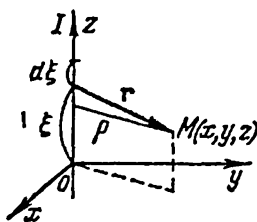


Fig. 224.

Let the conductor coincide with the z -axis, the current flowing upward (see Fig. 224), and let ξ denote the variable distance from the element of current to the origin, x, y and z being the coordinates of the point M . Then

$$d\mathbf{s} = d\xi \mathbf{k}, \quad \mathbf{r} = x\mathbf{i} + y\mathbf{j} + (z - \xi)\mathbf{k}$$

and

$$r = |\mathbf{r}| = \sqrt{x^2 + y^2 + (z - \xi)^2} = \sqrt{\rho^2 + (z - \xi)^2}$$

where $\rho\sqrt{x^2 + y^2}$ is the distance from the point M to the conductor. Computing the vector product we find $d\mathbf{H}$:

$$d\mathbf{H} = \frac{kI}{r^3} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 0 & d\xi \\ x & y & z - \xi \end{vmatrix} = \frac{kI}{r^3} (-y d\xi \mathbf{i} + x d\xi \mathbf{j})$$

This means that

$$dH_x = -\frac{kI y d\xi}{(\sqrt{\rho^2 + (z - \xi)^2})^3}, \quad dH_y = \frac{kI x d\xi}{(\sqrt{\rho^2 + (z - \xi)^2})^3} \quad \text{and} \quad dH_z = 0$$

To determine H_x and H_y we should integrate the corresponding differentials from $-\infty$ to ∞ . To this end let us evaluate the improper integral

$$\int_{-\infty}^{\infty} \frac{d\xi}{(\sqrt{\rho^2 + (z - \xi)^2})^3}$$

The substitution $\xi - z = \rho \tan t$, $d\xi = \frac{\rho}{\cos^2 t} dt$ leads to the integral

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\rho dt}{\cos^2 t (\sqrt{\rho^2 (1 + \tan^2 t)})^3} = \frac{1}{\rho^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos t dt = \frac{2}{\rho^2}$$

Thus,

$$H_x = -\frac{2kI y}{\rho^2} = -\frac{2kI y}{x^2 + y^2}, \quad H_y = \frac{2kI x}{\rho^2} = \frac{2kI x}{x^2 + y^2} \quad \text{and} \quad H_z = 0$$

This field is not defined at the points of the axis Oz . We see that the intensity vector $\mathbf{H}$ is in the same direction as the vector of linear velocity of a body rotating about Oz (see Sec. 151, I) if the direction of the electric current flow coincides with that of the angular velocity. The modulus of the vector $\mathbf{H}$ is

$$|\mathbf{H}| = \sqrt{H_x^2 + H_y^2} = \frac{2kl}{\rho}$$

It is readily checked that the divergence of the field $\mathbf{H}$ is equal to zero. Indeed, we have

$$\frac{\partial H_x}{\partial x} = \frac{4lxy}{(x^2 + y^2)^2}, \quad \frac{\partial H_y}{\partial y} = -\frac{4lxy}{(x^2 + y^2)^2} \quad \text{and} \quad \frac{\partial H_z}{\partial z} = 0$$

whence

$$\operatorname{div} \mathbf{H} = 0$$

The rotation of the field is also equal to zero at all points. To verify this property it is sufficient to show that

$$\frac{\partial H_x}{\partial y} - \frac{\partial H_y}{\partial x} = 0$$

the latter being readily seen:

$$\begin{aligned} \frac{\partial H_x}{\partial y} &= 2kl \frac{y^2 - x^2}{(x^2 + y^2)^2} \quad \text{and} \\ \frac{\partial H_y}{\partial x} &= 2kl \frac{y^2 - x^2}{(x^2 + y^2)^2} \end{aligned}$$

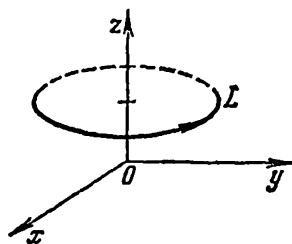


Fig. 225

Consequently, the circulation of the field over any contour *not encircling* Oz is equal to zero. But if we take a contour encircling the z -axis (see Fig. 225) this conclusion fails since any surface with such a contour as boundary intersects the axis Oz at whose points the field is not defined.

Let us compute the circulation round the circle of radius R with centre at the origin lying in the plane Oxy . This contour is represented by the equations

$$x = R \cos t, \quad y = R \sin t, \quad z = 0$$

Therefore we obtain

$$2kl \int_L \frac{-y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = 2kl \int_0^{2\pi} \frac{R^2 \sin^2 t + R^2 \cos^2 t}{R^2} dt = 4k\pi l$$

The circulation is independent of the radius R of the circle. It can be proved that it retains the same value for any closed contour encircling the z -axis.

157*. Nonstationary Fields. In this section we shall discuss some properties of *nonstationary* fields for which the specifying functions

are dependent not only on the spatial coordinates x, y and z but also on time t . A scalar field of this kind is determined by a function $u(x, y, z, t)$; the partial derivative $\frac{\partial u}{\partial t}$ serves as the time-rate of the magnitude u at the given fixed point $M(x, y, z)$.

A nonstationary vector field is specified by an equation

$$\mathbf{A}(x, y, z, t) = A_x(x, y, z, t)\mathbf{i} + A_y(x, y, z, t)\mathbf{j} + A_z(x, y, z, t)\mathbf{k}$$

In this case the derivative $\frac{\partial \mathbf{A}}{\partial t}$ is a vector given by the formula

$$\frac{\partial \mathbf{A}}{\partial t} = \frac{\partial A_x}{\partial t}\mathbf{i} + \frac{\partial A_y}{\partial t}\mathbf{j} + \frac{\partial A_z}{\partial t}\mathbf{k}$$

The derivatives $\frac{\partial u}{\partial t}$ and $\frac{\partial \mathbf{A}}{\partial t}$ are termed *partial time derivatives* or *explicit partial derivatives with respect to time t* ; they characterize the local rate of change of the fields u and $\mathbf{A}$ at the given point $M(x, y, z)$.

A nonstationary scalar field u generates the vector field of its gradient which is also nonstationary. Since $\text{grad } u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}$ we have

$$\frac{\partial (\text{grad } u)}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial x} \right) \mathbf{i} + \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial y} \right) \mathbf{j} + \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial z} \right) \mathbf{k}$$

On the other hand, the gradient of the scalar field $\frac{\partial u}{\partial t}$ is

$$\text{grad } \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial t} \right) \mathbf{i} + \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial t} \right) \mathbf{j} + \frac{\partial}{\partial z} \left(\frac{\partial u}{\partial t} \right) \mathbf{k}$$

The equality of mixed partial derivatives of the second order immediately implies that

$$\frac{\partial (\text{grad } u)}{\partial t} = \text{grad } \frac{\partial u}{\partial t}$$

Similarly, a nonstationary vector field generates the fields of its divergence and rotation which are also nonstationary (i.e. depend on t). It is easily shown that

$$\frac{\partial (\text{div } \mathbf{A})}{\partial t} = \text{div } \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \frac{\partial (\text{rot } \mathbf{A})}{\partial t} = \text{rot } \frac{\partial \mathbf{A}}{\partial t}$$

On both sides of the first formula there are scalar quantities while in the second formula we have vector quantities. In the derivation of these formulas we of course suppose that the conditions guaranteeing the equality of mixed partial derivatives of the second order are fulfilled for the differentiation with respect to the coordinates and with respect to time; by the theorem stated in Sec. 114,

the equality is guaranteed if these derivatives are continuous.

The formulas of vector analysis presented here play an important role in the study of nonstationary electromagnetic fields described by *Maxwell's*¹ equations.

Now let us discuss some problems arising in the study of processes in moving media. For instance, consider a distribution of temperature T in a fluid flow. In the nonstationary case the temperature depends both on the point where it is measured and on time: $T = T(x, y, z, t)$. If the temperature is measured at a fixed point $M(x, y, z)$ then, as was said, its time-rate is equal to the partial derivative $\frac{\partial T}{\partial t}$. But if we consider the temperature of a moving particle of the fluid it undergoes additional changes since the particle may pass from a warm part of the flow to a cold one and vice versa. Let the trajectory of motion of the particle be described by equations $x = x(t)$, $y = y(t)$, $z = z(t)$. Then the projections of the velocity $\mathbf{v}$ of this particle on the coordinate axes are

$$v_x = \frac{dx}{dt}, \quad v_y = \frac{dy}{dt} \quad \text{and} \quad v_z = \frac{dz}{dt}$$

The temperature $T(x, y, z, t)$ of the particle is a composite function of t , and its rate of change is equal to the *total partial derivative* with respect to t (see Sec. 116):

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{\partial T}{\partial x} v_x + \frac{\partial T}{\partial y} v_y + \frac{\partial T}{\partial z} v_z$$

The sum of the last three summands on the right-hand side is equal to the scalar product of grad T by the velocity $\mathbf{v}$; therefore

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \text{grad } T \cdot \mathbf{v}$$

Thus, the total partial derivative $\frac{dT}{dt}$ (also called the *particle derivative*) is the sum of the explicit partial derivative $\frac{\partial T}{\partial t}$ and the scalar product grad $T \cdot \mathbf{v}$. The latter describes the phenomenon of *convection* in the fluid and is called the *convective term* (the *convective derivative*).

QUESTIONS

1. How do we determine the work performed in the motion of a material point in a field of force?

2. State the definition of the line integral $\int_L P(x, y) dx + Q(x, y) dy$ over a given contour L .

¹ Maxwell, J.C. (1831-1879), a Scottish physicist.

are dependent not only on the spatial coordinates x, y and z but also on time t . A scalar field of this kind is determined by a function $u(x, y, z, t)$; the partial derivative $\frac{\partial u}{\partial t}$ serves as the time-rate of the magnitude u at the given fixed point $M(x, y, z)$.

A nonstationary vector field is specified by an equation

$$\mathbf{A}(x, y, z, t) = A_x(x, y, z, t)\mathbf{i} + A_y(x, y, z, t)\mathbf{j} + A_z(x, y, z, t)\mathbf{k}$$

In this case the derivative $\frac{\partial \mathbf{A}}{\partial t}$ is a vector given by the formula

$$\frac{\partial \mathbf{A}}{\partial t} = \frac{\partial A_x}{\partial t}\mathbf{i} + \frac{\partial A_y}{\partial t}\mathbf{j} + \frac{\partial A_z}{\partial t}\mathbf{k}$$

The derivatives $\frac{\partial u}{\partial t}$ and $\frac{\partial \mathbf{A}}{\partial t}$ are termed *partial time derivatives* or *explicit partial derivatives with respect to time t* ; they characterize the local rate of change of the fields u and $\mathbf{A}$ at the given point $M(x, y, z)$.

A nonstationary scalar field u generates the vector field of its gradient which is also nonstationary. Since $\text{grad } u = \frac{\partial u}{\partial x}\mathbf{i} + \frac{\partial u}{\partial y}\mathbf{j} + \frac{\partial u}{\partial z}\mathbf{k}$ we have

$$\frac{\partial (\text{grad } u)}{\partial t} = \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial x} \right) \mathbf{i} + \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial y} \right) \mathbf{j} + \frac{\partial}{\partial t} \left(\frac{\partial u}{\partial z} \right) \mathbf{k}$$

On the other hand, the gradient of the scalar field $\frac{\partial u}{\partial t}$ is

$$\text{grad } \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial t} \right) \mathbf{i} + \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial t} \right) \mathbf{j} + \frac{\partial}{\partial z} \left(\frac{\partial u}{\partial t} \right) \mathbf{k}$$

The equality of mixed partial derivatives of the second order immediately implies that

$$\frac{\partial (\text{grad } u)}{\partial t} = \text{grad } \frac{\partial u}{\partial t}$$

Similarly, a nonstationary vector field generates the fields of its divergence and rotation which are also nonstationary (i.e. depend on t). It is easily shown that

$$\frac{\partial (\text{div } \mathbf{A})}{\partial t} = \text{div } \frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \frac{\partial (\text{rot } \mathbf{A})}{\partial t} = \text{rot } \frac{\partial \mathbf{A}}{\partial t}$$

On both sides of the first formula there are scalar quantities while in the second formula we have vector quantities. In the derivation of these formulas we of course suppose that the conditions guaranteeing the equality of mixed partial derivatives of the second order are fulfilled for the differentiation with respect to the coordinates and with respect to time; by the theorem stated in Sec. 114,

the equality is guaranteed if these derivatives are continuous.

The formulas of vector analysis presented here play an important role in the study of nonstationary electromagnetic fields described by *Maxwell's¹ equations*.

Now let us discuss some problems arising in the study of processes in moving media. For instance, consider a distribution of temperature T in a fluid flow. In the nonstationary case the temperature depends both on the point where it is measured and on time: $T = T(x, y, z, t)$. If the temperature is measured at a fixed point $M(x, y, z)$ then, as was said, its time-rate is equal to the partial derivative $\frac{\partial T}{\partial t}$. But if we consider the temperature of a moving particle of the fluid it undergoes additional changes since the particle may pass from a warm part of the flow to a cold one and vice versa. Let the trajectory of motion of the particle be described by equations $x = x(t)$, $y = y(t)$, $z = z(t)$. Then the projections of the velocity $\mathbf{v}$ of this particle on the coordinate axes are

$$v_x = \frac{dx}{dt}, \quad v_y = \frac{dy}{dt} \quad \text{and} \quad v_z = \frac{dz}{dt}$$

The temperature $T(x, y, z, t)$ of the particle is a composite function of t , and its rate of change is equal to the *total partial derivative* with respect to t (see Sec. 116):

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \frac{\partial T}{\partial x} v_x + \frac{\partial T}{\partial y} v_y + \frac{\partial T}{\partial z} v_z$$

The sum of the last three summands on the right-hand side is equal to the scalar product of $\text{grad } T$ by the velocity $\mathbf{v}$; therefore

$$\frac{dT}{dt} = \frac{\partial T}{\partial t} + \text{grad } T \cdot \mathbf{v}$$

Thus, the total partial derivative $\frac{dT}{dt}$ (also called the *particle derivative*) is the sum of the explicit partial derivative $\frac{\partial T}{\partial t}$ and the scalar product $\text{grad } T \cdot \mathbf{v}$. The latter describes the phenomenon of *convection* in the fluid and is called the *convective term* (the *convective derivative*).

QUESTIONS

1. How do we determine the work performed in the motion of a material point in a field of force?

2. State the definition of the line integral $\int_L P(x, y, z) dx + Q(x, y, z) dy + R(x, y, z) dz$ over a given contour L .

¹ Maxwell, J.C. (1831-1879), a Scottish physicist.

3. How is the computation of the line integral reduced to the computation of the ordinary definite integral when the contour of integration is specified by parametric equations? Write down the computation formula.

4. How is the line integral evaluated when the contour of integration is represented by an equation of the form $y=y(x)$ or $x=x(y)$?

5. How does the change of the direction of integration affect the value of the line integral? How is the positive direction along a closed contour defined?

6. State and prove Green's theorem.

7. When do we say that the line integral is path-independent?

8. Show that path-independence of a line integral is equivalent to the condition that the integral is equal to zero for any closed contour.

9. State and prove the necessary and sufficient conditions for a line integral to be path-independent.

10. Which properties of the line integral and of its element of integration are equivalent in the case of a simply connected domain?

11. What is a primitive of a given total differential? Write down its general expression.

12. State the definition of a line integral over a space curve.

13. State and prove the path-independence theorem for the line integral over a spatial contour.

14. What is a potential field of force? What do we call a potential of such a field?

15. State the definition of a line integral with respect to arc length and enumerate the methods for its computation.

16*. What is a fluid flux across a given surface?

17*. What is a surface integral?

18*. What do we call a two-sided surface? Give an example of a one-sided surface.

19*. Give examples of problems leading to a surface integral whose integrand is not connected with the choice of the normal to the surface and also examples where there is such a connection. What do we mean by an integral taken over a definite side of a surface?

20*. How do we evaluate surface integrals?

21*. State and prove Stokes' theorem.

22*. Using Stokes' theorem prove the necessary and sufficient conditions for a line integral over a space curve to be path-independent.

23*. State and prove the Gauss-Ostrogradsky theorem.

24*. What is a vector field? Give concrete examples of vector fields.

- 25*. What do we call a vector line? Derive (differential) equations of vector lines.
- 26*. What is the flux of a vector field across a surface?
- 27*. State the definition of the divergence of a vector field. Derive the formula for the divergence.
- 28*. Write the Gauss-Ostrogradsky formula in vector form and explain its physical meaning.
- 29*. What is the circulation of a vector field?
- 30*. Derive the formula for the limit of the ratio of the circulation of a vector field over a plane contour L to the area bounded by the contour as the latter is contracted to a point.
- 31*. State the definition of the rotation of a vector field. Write Stoke's formula in vector form.
- 32*. Enumerate the operation rules for the Hamiltonian operator.
- 33*. Enumerate the vector field operations involving repeated application of the Hamiltonian operator.
- 34*. State the definition of a solenoidal field and enumerate its basic properties.
- 35*. Enumerate the basic properties of a potential field.
- 36*. What is a harmonic field? Write down the equation for the potential of such a field.

3. How is the computation of the line integral reduced to the computation of the ordinary definite integral when the contour of integration is specified by parametric equations? Write down the computation formula.

4. How is the line integral evaluated when the contour of integration is represented by an equation of the form $y=y(x)$ or $x=x(y)$?

5. How does the change of the direction of integration affect the value of the line integral? How is the positive direction along a closed contour defined?

6. State and prove Green's theorem.

7. When do we say that the line integral is path-independent?

8. Show that path-independence of a line integral is equivalent to the condition that the integral is equal to zero for any closed contour.

9. State and prove the necessary and sufficient conditions for a line integral to be path-independent.

10. Which properties of the line integral and of its element of integration are equivalent in the case of a simply connected domain?

11. What is a primitive of a given total differential? Write down its general expression.

12. State the definition of a line integral over a space curve.

13. State and prove the path-independence theorem for the line integral over a spatial contour.

14. What is a potential field of force? What do we call a potential of such a field?

15. State the definition of a line integral with respect to arc length and enumerate the methods for its computation.

16*. What is a fluid flux across a given surface?

17*. What is a surface integral?

18*. What do we call a two-sided surface? Give an example of a one-sided surface.

19*. Give examples of problems leading to a surface integral whose integrand is not connected with the choice of the normal to the surface and also examples where there is such a connection. What do we mean by an integral taken over a definite side of a surface?

20*. How do we evaluate surface integrals?

21*. State and prove Stokes' theorem.

22*. Using Stokes' theorem prove the necessary and sufficient conditions for a line integral over a space curve to be path-independent.

23*. State and prove the Gauss-Ostrogradsky theorem.

24*. What is a vector field? Give concrete examples of vector fields.

- 25*. What do we call a vector line? Derive (differential) equations of vector lines.
- 26*. What is the flux of a vector field across a surface?
- 27*. State the definition of the divergence of a vector field. Derive the formula for the divergence.
- 28*. Write the Gauss-Ostrogradsky formula in vector form and explain its physical meaning.
- 29*. What is the circulation of a vector field?
- 30*. Derive the formula for the limit of the ratio of the circulation of a vector field over a plane contour L to the area bounded by the contour as the latter is contracted to a point.
- 31*. State the definition of the rotation of a vector field. Write Stoke's formula in vector form.
- 32*. Enumerate the operation rules for the Hamiltonian operator.
- 33*. Enumerate the vector field operations involving repeated application of the Hamiltonian operator.
- 34*. State the definition of a solenoidal field and enumerate its basic properties.
- 35*. Enumerate the basic properties of a potential field.
- 36*. What is a harmonic field? Write down the equation for the potential of such a field.

Chapter X

DIFFERENTIAL EQUATIONS

§ 1. Differential Equations of the First Order

158. General Notions. Existence Theorem. When studying the integral calculus of functions of one independent variable we dealt with the problem of finding an unknown function y from its derivative or differential.

The equation

$$y' = f(x) \text{ or, equivalently, } dy = f(x) dx \quad (*)$$

where y is an unknown function of x and $f(x)$ a given continuous function, is the simplest *differential equation*. To solve the equation, i.e. to determine the unknown function y , we must integrate the given function $f(x)$. As is known, this leads to an infinitude of functions each of which satisfies condition (*). In this chapter it will be convenient to understand the symbol $\int f(x) dx$ as designating a *concrete antiderivative* of $f(x)$. Then any solution of equation (*) is written in the form

$$y = \int f(x) dx + C$$

As will be seen, we usually deal with equations of a more complicated type which contain, besides the derivative y' and the independent variable x , the unknown function y itself. Examples of such equations are

$$y' + x^2y = 0, \quad xy' - y = 0, \quad xy' = y + x, \text{ etc.}$$

Replacing y' by $\frac{dy}{dx}$ we can rewrite the same equations in the form involving the differentials instead of the derivatives:

$$dy + x^2y dx = 0, \quad x dy - y dx = 0, \quad x dy - (y + x) dx = 0$$

Definition. A *differential equation of the first order* is an equation connecting the independent variable, the unknown function and its derivative.

The derivative being representable as the ratio of the differentials, the equation may involve the differentials of the unknown function and of the independent variable instead of the derivative.

Differential equations of the *second* and *higher* orders will be considered in § 2.

In this course we shall limit ourselves to studying equations whose unknown functions depend solely on *one* argument¹.

The general form of a first-order differential equation is

$$F(x, y, y') = 0$$

In particular cases the left-hand side of the equation² may not contain x or y but the derivative y' always enters into it. We shall mainly deal with equations resolved with respect to the derivative which are written in the form

$$y' = f(x, y)$$

Definition. Every function which, when substituted, together with its derivative, into the given differential equation, turns the equation into an identity is called a *solution of the differential equation*.

The simplest examples show that, generally, a differential equation possesses an infinite number of solutions. This is, for instance, clearly demonstrated by equation (*). It is also readily checked that the solutions of the equation $y' = y$ are the functions $y = Ce^x$ while the equation $y' = -y$ is satisfied by the functions $y = Ce^{-x}$ where C is any constant number.

Consider the equation $y' = \frac{y+x}{x}$. For $x > 0$ its solutions are the functions $y = x \ln x + Cx$. For, on finding the derivative $y' = \ln x + 1 + C$ and substituting it into the equation, we obtain the identity

$$\ln x + 1 + C \equiv \frac{x \ln x + Cx + x}{x}$$

It is easily verified that for $x < 0$ the solutions are the functions $y = x \ln |x| + Cx$.

We see that the solutions of the above differential equations contain an *arbitrary constant* C . Now, although all these equations are particular examples, we draw the following general conclusion:

¹ These equations are known as *ordinary differential equations*. The differential equations involving unknown functions of several arguments are called *partial differential equations* and are treated in courses devoted to *equations of mathematical physics*.

² Unless a confusion may occur, we shall often say "equation" instead of "differential equation".

The differential equation of the first order $y' = f(x, y)$ possesses in the general case, an infinitude of solutions which are usually expressed by a formula $y = \varphi(x, C)$ containing one arbitrary constant (it may happen that C can only vary within some limits).

Such a family of solutions is called the *general solution*. Making the arbitrary constant C assume some admissible concrete numerical values, we obtain *particular solutions*.

In what follows, when solving concrete problems, we shall most often be concerned with particular solutions specified by the so-called *initial conditions* which are set in order to isolate the solutions we are interested in.

An *initial condition* for a first-order differential equation is usually introduced by setting a pair of values x_0 and y_0 of the independent variable x and of the unknown function y which correspond to each other. This condition is written as

$$y|_{x=x_0} = y_0$$

Now it is natural to pose the question as to whether there exists a solution of the equation $y' = f(x, y)$ satisfying the given initial condition $y|_{x=x_0} = y_0$ and, provided it exists, whether this solution is unique. The answer to the question is given by the following theorem which was stated and proved by Cauchy (for this reason the problem of determining the particular solution specified by a given initial condition is referred to as *Cauchy's problem*):

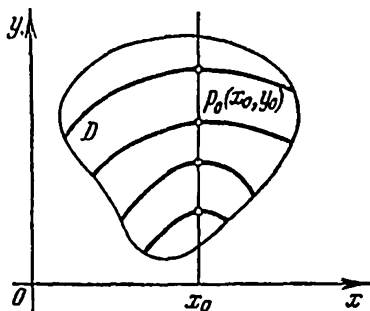


Fig. 226

Cauchy's Theorem on the Existence and Uniqueness of the Solution. Given an equation $y' = f(x, y)$ with an initial condition $y|_{x=x_0} = y_0$,

the particular solution $y = y(x)$ satisfying this initial condition (i.e. such that $y(x_0) = y_0$) exists in a sufficiently small interval $[x_0 - h, x_0 + h]$ and is unique provided that the function $f(x, y)$ and its partial derivative $\frac{\partial f}{\partial y}$ are continuous in an open region containing the point $P_0(x_0, y_0)$.

It should be noted that the existence of the solution is only guaranteed for a sufficiently small interval $[x_0 - h, x_0 + h]$ and the uniqueness is guaranteed within the open domain under consideration.

In our course we shall not present the proof of this theorem.

Suppose that the conditions of Cauchy's theorem are fulfilled. Then in an interval enclosing the point x_0 there exists a unique solution $y = y(x)$ whose graph passes through the point $P_0(x_0, y_0)$.

If x_0 is fixed and y_0 is varied and the conditions of Cauchy's theorem hold in a neighbourhood of every point (x_0, y_0) we obtain different particular solutions for different y_0 (see Fig. 226). This demonstrates visually that the differential equation $y' = f(x, y)$ has infinitely many solutions. These solutions depend on the choice of the value $y_0 = C$; if $y_0 = C$ is taken as the arbitrary constant we can write $y = \varphi(x, C) = \varphi(x, y_0)$.

Usually the graphs of the solutions entirely fill a region in the xy -plane and cover it in such a way that only a single curve passes through each point of that region. Then the expression $y = \varphi(x, C)$ describes the general solution in this region.

The graph of a particular solution of a differential equation is called an *integral curve* of the equation. To the general solution $y = \varphi(x, C)$ there corresponds a *family of integral curves*.

The introduction of an initial condition $y|_{x=x_0} = y_0$ is equivalent to setting a point $P_0(x_0, y_0)$ through which the integral curve corresponding to the sought-for particular solution must pass.

Hence, given an initial condition, to find the particular solution corresponding to it, we must choose from the family of integral curves the one passing through the point $P_0(x_0, y_0)$. If the general solution $y = \varphi(x, C)$ of the differential equation is known we simply substitute x_0 and y_0 into it and thus obtain the equation $y_0 = \varphi(x_0, C)$ for determining C .

Examples. (1) The equation $y' = 2x$ has the general solution $y = x^2 + C$ which represents a family of parabolas (see Fig. 227). The right-hand side of this equation obviously satisfies the conditions of Cauchy's theorem in the whole plane, and therefore only a single integral curve passes through each point of the plane.

To find the particular solution satisfying the initial condition $y|_{x=1} = 2$ we substitute the given values into the general solution: $2 = 1^2 + C$. Whence $C = 1$. Consequently, the sought-for particular solution is $y = x^2 + 1$.

(2) As was shown, the equation $y' = y$ has the general solution $y = Ce^x$. The corresponding family of integral curves (the collection of the graphs of exponential functions) is depicted in Fig. 228; it also includes the axis of abscissas (for $C = 0$). In this example the conditions of Cauchy's theorem also hold in the whole xy -plane.

Let us determine the particular solution satisfying the initial condition $y|_{x=0} = 2$. The substitution of these data into the general solution yields $C = 2$. Hence, the function $y = 2e^x$ is the particular solution we have been looking for.

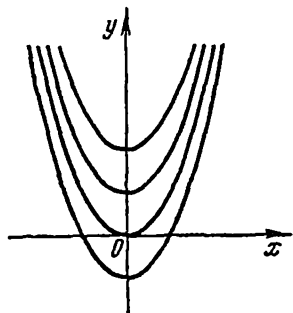


Fig. 227

(3) For the equation $y' = \frac{x+y}{x}$ the conditions of Cauchy's theorem are, for instance, fulfilled in the half-plane $x > 0$. As was shown, in this region the solutions are given by the formula $y = x \ln x + Cx$; this is the general solution in the domain under consideration. In this example the family of integral curves is of a more complicated structure.

The general solution of a differential equation is by far not always expressible in an explicit form $y = \varphi(x, C)$, and it is sometimes necessary to represent it as an implicit function: $\Phi(x, y, C) = 0$; such a relation is usually called the *general integral* of the differential equation. Generally, the process of finding solutions of a differential equation is referred to as *integration of the equation*. This terminology by no means implies that the process is reducible to the ordinary integration of functions, which is only the case for some simpler types of equation.

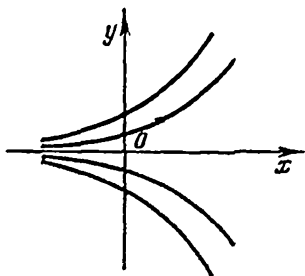


Fig. 228

159. Equations with Variables Separable. Let us consider an equation of the form

$$f_1(y) dy = f_2(x) dx \quad (*)$$

where $f_1(y)$ and $f_2(x)$ are given functions. In this differential equation the variables are *separated* in the sense that each variable

only enters into the member of the equation containing its differential. The equation $dy = f(x) dx$ is a particular case of equations (*).

Both members of equation (*) are differentials of certain functions; the differential on the right-hand side is directly expressed in terms of the independent variable x while the one on the left involves the intermediate argument y which is a function of x . It is this functional relationship between y and x that must be found in this problem. Performing integration we obtain a new relationship, connecting the variables x and y , which no longer involves the differentials (since the signs of the integral and of the differential mutually cancel out):

$$\int f_1(y) dy = \int f_2(x) dx + C$$

Remember that in this chapter we agreed to designate only one concrete antiderivative by the symbol $\int$. It is clear that the arbitrary constant C can also be written on the right-hand side of the last equality.

If we are given an initial condition $y|_{x=x_0} = y_0$ the particular solution satisfying this condition is found after the corresponding value of C has been determined. Using definite integrals with

variable limits of integration we can immediately write down the relation for the sought-for particular solution:

$$\int_{y_0}^y f_1(y) dy = \int_{x_0}^x f_2(x) dx$$

Here the values x_0 and y_0 do in fact correspond to each other since both members of the equality vanish when the upper limits y and x are, respectively, replaced by y_0 and x_0 .

Thus, the solution of differential equation (*) is written in the form containing integrals of given functions; in such cases we say that the equation is *reducible to quadratures* (*integrable by quadratures*).

Examples. (1) Let us find the general solution of the equation $3y^2 dy = 2x dx$. On integrating we get $y^3 = x^2 + C$; in this case the solution is readily found in the explicit form: $y = \sqrt[3]{x^2 + C}$.

(2) Consider the equation $\frac{dy}{y} = 3x^2 dx$. Let us find its solution satisfying the condition $y|_{x=0} = 2$.

On integrating and writing the arbitrary constant in the form $\ln|C|$ ($C \neq 0$), which is convenient for taking antilogarithms, we obtain $\ln|y| = x^3 + \ln|C|$, whence

$$y = Ce^{x^3}$$

If we took the arbitrary constant in the form $\ln C$ the general solution would be written as $y = \pm Ce^{x^3}$ because in this case we must have $C > 0$.

Substituting the given initial values into the general solution we determine C :

$$2 = C$$

Thus, the function $y = 2e^{x^3}$ is the sought-for particular solution of the given equation. We could also have used definite integrals, which would lead to

$$\int_2^y \frac{dy}{y} = \int_0^x 3x^2 dx$$

that is,

$$\ln y \Big|_2^y = x^3 \Big|_0^x, \text{ or } \ln y - \ln 2 = x^3$$

Now, taking antilogarithms we arrive at the same particular solution:

$$y = 2e^{x^3}$$

We also often deal with equation in which the variables are not separated originally but can be separated by means of simple arithmetical operations.

Definition. A differential equation in which the variables can be separated with the aid of the multiplication or division of both members of the equation by one and the same expression is called a *differential equation with variables separable*.

For instance, such is the equation

$$\frac{dy}{dx} = \frac{f_2(x)}{f_1(y)}$$

Here the variables are not yet separated but we obtain an equation with separated variables on multiplying both sides by $f_1(y)dx$.

The variables are also readily separated if the equation has the form

$$f_1(x) f_2(y) dx + f_3(x) f_4(y) dy = 0 \quad (**)$$

involving the differentials dx and dy .

On dividing both sides of the equation by the product $f_2(y) f_3(x)$ we obtain

$$\frac{f_1(x)}{f_3(x)} dx + \frac{f_4(y)}{f_2(y)} dy = 0$$

Here it is unnecessary to transpose one of the summands to the right-hand side. Integrating we get

$$\int \frac{f_1(x)}{f_3(x)} dx + \int \frac{f_4(y)}{f_2(y)} dy = C$$

It should be noted that the division by $f_2(y) f_3(x)$ may lead to the omission of some particular solutions of equations (**). For instance, let $f_2(y_0)$ vanish for $y = y_0$: $f_2(y_0) = 0$. Then the constant function $y = y_0$ is a solution of the equation. Indeed, we have $dy = 0$, and hence the substitution of $y = y_0$ into equation (**) leads to an identity. If x and y are regarded as playing equivalent roles in the equation, analogous argument shows that if $f_3(x_0) = 0$, then $x = x_0$ is also a particular solution of the equation¹.

(3) Let us find the general solution of the equation

$$x(y+1)dx - (x^2+1)ydy = 0$$

Separating the variables we bring it to the form

$$\frac{x dx}{x^2+1} - \frac{y dy}{y+1} = 0$$

Now the integration yields

$$\frac{1}{2} \ln(x^2+1) - \int \left(1 - \frac{1}{y+1}\right) dy = C$$

¹ In the answers given in collections of problems such solutions are usually omitted.

i.e.

$$\frac{1}{2} \ln(x^2 + 1) - y + \ln|y + 1| = C$$

Thus, we have found the general solution, the variable y being an implicit function of x . Replacing C by $\ln|C|$, we can write the solution in the form

$$\frac{Ce^y}{y+1} = \sqrt{x^2 + 1}$$

In addition, here there exists one more particular solution: $y = -1$ (its graph is the horizontal line $y = -1$).

160. Some Physical Problems. Although the differential equations with variables separable are a narrow particular class of first-order equations, there are many problems of physics and mechanics leading to them.

It should be noted that, as a rule, the main difficulty in solving a problem of physical character with the aid of differential equations lies in forming the differential equation itself. There exist no general rules for setting such an equation, and each problem requires an individual approach. We shall consider some typical examples of this kind to demonstrate techniques of setting differential equations. These examples will help the reader to understand the essence of these approaches and proceed to solving similar problems.

(1) **Radioactive decay.** It was established experimentally that *the rate of radioactive decay is proportional to the remaining (not disintegrated) mass of the substance*. Let the initial mass of the radioactive substance be M_0 . We shall determine the relationship between the amount M of the remaining substance and time t .

The rate of radioactive decay is equal to the derivative of the mass M with respect to time t , that is, to $\frac{dM}{dt}$. Now, according to the above law, we can write the relation

$$\frac{dM}{dt} = -kM$$

where $k > 0$ is a proportionality coefficient. It is taken with the minus sign since the amount of the substance M decreases as t grows, which indicates that the derivative is nonpositive. Separating the variables in the equation thus obtained we write

$$\frac{dM}{M} = -k dt$$

On integrating we obtain

$$\ln M = -kt + \ln C$$

whence

$$M = Ce^{-kt} \quad (*)$$

Note that the quantity M_0 does not enter into the differential equation; it only appears in the initial condition which has the form $M|_{t=0} = M_0$. This condition implies that $C = M_0$. Consequently, the particular solution satisfying the conditions of the problem is

$$M = M_0 e^{-kt}$$

The value of the constant k can be determined experimentally by measuring the amount of the remaining substance at a time moment t .¹

The rate of decay is usually characterized by the so-called *half-life period* (or, simply, *half-life*) T of the substance which is the time during which the initial value of M becomes half as large.

It follows from formula (*) that $\frac{M_0}{2} = M_0 e^{-kT}$ and, hence, $e^{kT} = 2$ and $T = \frac{\ln 2}{k}$. It is clearly senseless to speak of the time of decay of the entire mass M_0 since, theoretically, it is infinite.

(2) *Cooling of a body.* According to the law established by Newton, *the rate of cooling of a physical body is directly proportional to the difference between the temperature of the body and that of the surrounding medium.*

Let at time moment $t = t_0 = 0$ the temperature of the body be T_0 . We assume that the temperature of the medium is constant² and equal to T_c ($T_c < T_0$). We shall be concerned with the problem of determining the relationship between the variable temperature of the body and time t .

Let the temperature of the body at time moment t be T . By Newton's law the rate of change of the temperature, i.e. $\frac{dT}{dt}$, is proportional to the difference $T - T_c$. Consequently,

$$\frac{dT}{dt} = -k(T - T_c)$$

We assume that $k > 0$ and take the minus sign on the right side of this relation since as time t increases the temperature T of the body decreases. The proportionality factor k depends both on the physical properties of the body and on its geometric form (for instance, it is apparent that the rate of cooling of a laminated steel plate is greater than that of the original ingot).

Separating the variables we obtain $\frac{dT}{T - T_c} = -k dt$, whence

$$\ln(T - T_c) = -kt + \ln C \quad \text{and} \quad T = T_c + Ce^{-kt}$$

¹ This method of determining the value of a physical constant is often used in many similar cases.

² If the temperature of the surrounding medium is variable (for instance, it may increase as the body is cooled) the problem becomes more complicated.

Since $T > T_c$ we can write $\ln(T - T_c)$ instead of $\ln|T - T_c|$. Let the reader consider the case $T_c > T_0$ when the body is not cooled but heated.

Substituting the initial values given by the initial condition $T|_{t=0} = T_0$ we find C : $T_0 = T_c + C$ whence $C = T_0 - T_c$. Finally, the law of cooling is written in the form

$$T = T_c + (T_0 - T_c)e^{-kt}$$

Here the proportionality factor k is regarded as being known. Its value can be determined experimentally by measuring the temperature T at a time moment t .¹ Theoretically, the temperature of the body only becomes equal to that of the surrounding medium as $t \rightarrow \infty$.

In both examples the corresponding physical law directly implies the relationship between the derivative of the unknown function and the function itself (the independent variable t does not enter into the equation).

We now consider a more complicated example.

(3) **Outflow of liquid from a cylindrical vessel.** Consider a cylindrical vessel filled with a liquid and having an opening at the bottom. Let us determine the time of the outflow of the liquid from the vessel if the initial height of the column of the liquid is H , the radius of the cylinder r and the area of the opening s (Fig. 229).

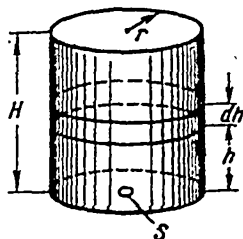


Fig. 229

The solution of the problem is based on *Toricelli's law*² which states that if the dimensions of the opening are small compared with the height of the column of the liquid the rate of outflow v is equal to $\sqrt{2gh}$ where h is the height of the liquid above the opening and g the acceleration of gravity³. (This rate is equal to the velocity with which a body falling freely from a height h reaches the ground.)

If the loss of the liquid were permanently compensated, the rate of the outflow specified by Toricelli's law would remain constant and the original amount of the liquid would disappear from the

¹ For instance, if it is known that $T = T_1$ for $t = t_1$ then $k = \frac{1}{t_1} \ln \frac{T_0 - T_c}{T_1 - T_c}$.

² E. Toricelli (1608-1647), a prominent Italian mathematician and physicist, a disciple and collaborator of the great Italian scientist Galileo Galilei.

³ Toricelli's law thus formulated is only applicable to an ideal liquid. In practice a more precise formula is used: $v = \mu \sqrt{2gh}$ where μ is a coefficient dependent on the viscosity of the liquid and the shape of the outflow opening. For instance, for water outflowing from a circular opening $\mu \approx 0.6$.

vessel in time $\frac{\pi r^2 H}{\sqrt{2gH}s}$. But without such a compensation the height of the liquid is constantly decreasing, and the problem becomes more complicated.

Let us denote the height of the liquid in the cylinder at time moment t by h (see Fig. 229). The connection between dt and dh can be found if we suppose that during an infinitesimal time interval dt the rate of outflow remains constant and is equal to $\sqrt{2gh}$. Then the volume dV of the liquid flowing out of the vessel in time dt is equal to the volume of a cylinder with base s and altitude vdt :

$$dV = s\sqrt{2gh} dt$$

Let the reader show that this volume differs from the exact volume of the outflowed liquid by an infinitesimal of higher order of smallness relative to dt .

On the other hand, since the height of the liquid decreases by dh the same volume is expressed as

$$dV = -\pi r^2 dh$$

(the minus sign has been put here since $dh < 0$). Equating the two expressions for dV we arrive at the differential equation

$$s\sqrt{2gh} dt = -\pi r^2 dh$$

Let us rewrite the equation in the form

$$dt = -\frac{\pi r^2}{s\sqrt{2g}h} dh$$

The integration of the latter equation results in

$$t = -\frac{2\pi r^2}{s\sqrt{2g}}\sqrt{h} + C$$

Since $h|_{t=0} = H$, we have

$$C = \frac{2\pi r^2}{s\sqrt{2g}}\sqrt{H}$$

Finally,

$$t = \frac{2\pi r^2}{s\sqrt{2g}}(\sqrt{H} - \sqrt{h})$$

The resultant formula expresses the time of outflow t as a function of the height of the level of the liquid h . Putting $h=0$ we determine the total time of outflow:

$$T = \frac{2\pi r^2}{s\sqrt{2g}}\sqrt{H}$$

161. Homogeneous Equations of the First Order. Linear Equations. In this section we shall consider some important classes of

simple equations of the first order reducible to equations with variables separable.

1. Homogeneous equations.

Definition. The equation

$$y' = f(x, y)$$

is said to be *homogeneous* if the function $f(x, y)$ can be represented as being dependent solely on the ratio of its arguments:

$$f(x, y) = \varphi\left(\frac{y}{x}\right)$$

For example, the equation

$$(xy - y^2) dx - (x^2 - 2xy) dy = 0$$

is homogeneous since it can be rewritten in the form

$$\frac{dy}{dx} = \frac{xy - y^2}{x^2 - 2xy} = \frac{\frac{y}{x} - \left(\frac{y}{x}\right)^2}{1 - 2\frac{y}{x}}$$

In the general case the variables entering into a homogeneous equation are not separated, but the introduction of an auxiliary function $u(x)$ determined by the formula

$$\frac{y}{x} = u \quad \text{or} \quad y = xu$$

transforms the homogeneous equation to an equation with variables separable.

Indeed, the derivative y' is expressed as

$$y' = u + xu'$$

and hence the equation $y' = \varphi\left(\frac{y}{x}\right)$ takes the form

$$u + xu' = \varphi(u), \quad \text{that is, } x \frac{du}{dx} = \varphi(u) - u$$

It follows that

$$\frac{du}{\varphi(u) - u} = \frac{dx}{x}$$

and after the integration we obtain

$$\int \frac{du}{\varphi(u) - u} = \ln|x| + C$$

On finding from this relation the variable u as function of x and returning to the original unknown function $y = xu$ we receive the sought-for solution of the homogeneous equation.

In most cases it is impossible to find an explicit expression for u . Then, after the integration, we should replace u in the left-hand side by $\frac{y}{x}$, which leads to the solution of the equation in implicit form.

It is of course supposed that $\varphi(u) - u \neq 0$. If $\varphi(u) \equiv u$ then $\varphi\left(\frac{y}{x}\right) \equiv \frac{y}{x}$, and no transformation is needed since the given equation $y' = \frac{y}{x}$ is one with the variables separable.

As has been already pointed out, if the denominator $\varphi(u) - u$ turns into zero only for an isolated value u_0 , the function $u = u_0$ is a solution of the transformed equation and the function $y = u_0 x$ satisfies the original equation.

It is unnecessary to learn by heart the above formulas since it is simpler to carry out all the calculations in every concrete case.

Example. Let us solve the homogeneous equation

$$y' = \frac{xy - y^2}{x^2 - 2xy}$$

The substitution $y = xu$ leads to the equation

$$u + xu' = \frac{u - u^2}{1 - 2u} \text{ or, equivalently, } \frac{du}{dx} = \frac{1}{x} \left(\frac{u - u^2}{1 - 2u} - u \right) = \frac{1}{x} \frac{u^3}{1 - 2u}$$

Separating the variables we receive

$$\frac{1 - 2u}{u^3} du = \frac{dx}{x}$$

whence $\frac{1}{u} + 2 \ln |u| = \ln \left| \frac{C}{x} \right|$. Hence, $\ln \left(e^{\frac{1}{u}} u^2 \right) = \ln \left| \frac{C}{x} \right|$, and consequently,

$$u^2 e^{\frac{1}{u}} = \frac{C}{x}$$

On returning to the variable y we arrive at the general solution

$$\frac{y^2}{x} e^{\frac{x}{y}} = C$$

II. Linear equations. Another frequently encountered type of first-order equations are linear equations.

Definition. An equation of the form

$$y' + p(x)y = q(x) \quad (*)$$

which is linear with respect to the unknown function and its derivative is called **linear**. Here $p(x)$ and $q(x)$ are given continuous functions of the independent variable x .

The following artificial technique reduces equation (*) to two equations with variables separable. Let us represent the function y as a product of two functions: $y = u(x)v(x)$. One of these functions can be taken quite arbitrarily and the other should then be determined, depending on the former, in such a way that the whole product satisfies the given linear equation. The arbitrariness in the choice of one of the functions u and v is used for simplifying the resultant equation appearing after the substitution.

From the equality $y = uv$ we find the derivative y' :

$$y' = u'v + uv'$$

Substituting this expression into equation (*) we obtain

$$u'v + uv' + p(x)uv = q(x), \text{ i.e. } u'v + u(v' + p(x)v) = q(x)$$

Now let us take as v an *arbitrary* particular solution of the equation

$$v' + p(x)v = 0 \quad (**)$$

Then we obtain the equation

$$u'v = q(x) \quad (***)$$

for determining u .

We shall begin with finding v from equation (**). On separating the variables we obtain

$$\frac{dv}{v} = -p(x) dx$$

whence

$$\ln v = - \int p(x) dx, \text{ that is, } v = e^{- \int p(x) dx}$$

As was agreed, the sign of the indefinite integral designates here a particular antiderivative of $p(x)$, and therefore v is a completely determined function of x .

Now, knowing v , we readily find u from equation (***):

$$\frac{du}{dx} = \frac{q(x)}{v} = q(x)e^{\int p(x) dx}, \quad du = q(x)e^{\int p(x) dx} dx$$

and, hence,

$$u = \int q(x)e^{\int p(x) dx} dx + C$$

In the latter formula expressing u we take the family of all the antiderivatives. Finally, knowing u and v we determine the sought-for function y :

$$y = uv = e^{- \int p(x) dx} \left[\int q(x)e^{\int p(x) dx} dx + C \right]$$

This formula expresses the *general solution* of linear equation (*).

The final result remains unchanged if we add an arbitrary constant C_1 to the integrals standing as exponents. Indeed, this second arbitrary constant is cancelled out in the ultimate formula since one factor has e^{C_1} in the denominator while the other has it in the numerator.

The problem can also be solved with the aid of definite integrals with variable limits of integration. This yields

$$v = e^{-\int_{x_0}^x p(x) dx}, \quad u = \int_{x_0}^x q(x) e^{\int_{x_0}^x p(x) dx} dx + C$$

and

$$y = e^{-\int_{x_0}^x p(x) dx} \left(\int_{x_0}^x q(x) e^{\int_{x_0}^x p(x) dx} dx + C \right)$$

The particular solution corresponding to the initial condition $y|_{x=x_0} = y_0$ is obtained from the last formula if we put $C = y_0$.

Here again we underline that instead of learning by heart the general formula it is preferable to memorize the method of solution and apply it to every concrete problem.

Examples. (1) Let us solve the equation $y' + \frac{1}{x}y = \frac{\sin x}{x}$.

We put $y = uv$; then $y' = u'v + uv'$. Furthermore, $u'v + uv' + \frac{1}{x}uv = \frac{\sin x}{x}$, and thus $u'v + u\left(v' + \frac{v}{x}\right) = \frac{\sin x}{x}$. Let $v' + \frac{v}{x} = 0$. Then $\frac{dv}{v} = -\frac{dx}{x}$, and, hence, $\ln|v| = -\ln|x|$, i.e. $v = \frac{1}{x}$. Consequently, $u' \cdot \frac{1}{x} = \frac{\sin x}{x}$, whence $u' = \sin x$ and $u = -\cos x + C$. Finally, we obtain

$$y = uv = \frac{1}{x}(-\cos x + C)$$

(2) Let us find the solution of the equation $y' - xy = 1$ satisfying the initial condition $y|_{x=0} = 0$. Here $p(x) = -x$, $q(x) = 1$, $x_0 = 0$ and $y_0 = 0$. By the general formula we obtain

$$y = e^{\frac{x^2}{2}} \int_0^x e^{-\frac{x^2}{2}} dx$$

Despite the fact that the integral on the right-hand side cannot be expressed in terms of elementary functions (see Sec. 85) the problem is considered as being solved since an expression for the function y involving quadratures has been found. The values of y can be determined by means of numerical integration. By the

way, in this case it is unnecessary to resort to approximate integration since there are extensive tables of values for the integral $\int_0^x e^{-\frac{x^2}{2}} dx$ (more precisely, the tables are compiled for the function

$\Phi(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$ known as the *probability integral* or *Laplace's function*).

There are some more complicated equations which can be reduced to linear ones. For instance, let us take the so-called *Bernoulli's¹ equation*

$$y' + p(x)y = q(x)y^n$$

For $n=0$ it becomes a linear equation and for $n=1$ its variables are separable. For other values of n it is reduced to a linear equation with the help of the following artificial transformation. We divide both members of the equation by y^n and rewrite it in the form

$$y^{-n}y' + p(x)y^{-n+1} = q(x)$$

The introduction of the auxiliary function $y^{-n+1} = z$ gives $(-n+1)y^{-n}y' = z'$, which transforms the equation to

$$z' + (-n+1)p(x)z = (-n+1)q(x)$$

The latter is a linear equation. On solving this equation and returning from z to y we obtain the solution of the original equation.

Now we proceed to an important electrical engineering problem leading to a linear equation of the first order.

Consider an electric circuit containing a resistance R , an inductance L and an electric-current source with electromotive force $\mathcal{E}$ (see Fig. 230a). As is known from physics, if I is the electric current flow then

$$\mathcal{E} = RI + L \frac{dI}{dt}$$

This is a linear differential equation with respect to the unknown function $I = I(t)$ which can be written in the form

$$\frac{dI}{dt} + \frac{R}{L}I = \frac{\mathcal{E}}{L}$$

We shall solve this equation with the initial condition $I|_{t=0} = 0$. Thus, we are concerned with the problem of *switching on* an electric current source.

¹ Jacob Bernoulli (1654-1705) and Johann Bernoulli (1667-1748), Swiss mathematicians, contemporaries of Leibniz and Newton, whose extensive scientific research and pedagogical activity contributed much to the development of mathematical analysis.

Making use of the general formula expressing the solution of a linear equation in terms of definite integrals we obtain

$$I = e^{-\int_0^t \frac{R}{L} dt} \left(\int_0^t \frac{\mathcal{E}}{L} e^{\int_0^t \frac{R}{L} dt} dt \right)$$

Performing the integration we get

$$I = \frac{\mathcal{E}}{R} \left(1 - e^{-\frac{R}{L} t} \right)$$

This formula describes the process of growth of the current in the circuit after the electric source has been switched on; the graph representing the dependence of I on t is shown in Fig. 230b

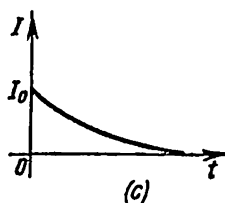
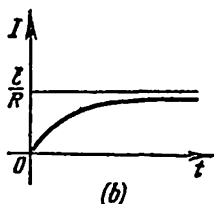
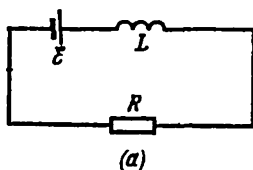


Fig. 230

Now let us proceed to the problem of *switching off* the electric current source (for the same circuit). In this case we put $I_{t=0} = I_0$ and $\mathcal{E} = 0$.¹ This leads to the equation with variables separable

$\frac{dI}{dt} = -\frac{R}{L} I$. Solving the equation we obtain $I = I_0 e^{-\frac{R}{L} t}$, which shows that after the source has been switched off the current does not vanish instantaneously but decreases in accordance with an exponential law; the graph of this function is shown in Fig. 230c.

We suggest that the reader consider the case when a sinusoidal voltage $U = U_0 \sin \omega t$ is applied to the terminals of the circuit.

¹ Here we do not discuss the question of technical realization of such a circuit.

162. Exact Differential Equations. Let us write down the differential equation of the first order in the general form involving differentials:

$$P(x, y)dx + Q(x, y)dy = 0 \quad (A)$$

Definition. If the left-hand side of equation (A) is the total differential of a function $u(x, y)$ it is called an *exact differential equation*.

If the functions $P(x, y)$ and $Q(x, y)$ and their partial derivatives are continuous in a simply connected domain equation (A) is exact if and only if $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.

Using the methods of reconstructing a function from its total differential (see Sec. 115, I and Sec. 142) we find a function $u(x, y)$ such that

$$du(x, y) = P(x, y)dx + Q(x, y)dy$$

Then equation (A) can be written as

$$du(x, y) = 0$$

The last equality means that the variables x and y are connected by the functional relationship

$$u(x, y) = C \quad (*)$$

where C is an arbitrary constant. This relationship expresses the general solution of equation (A). Consequently, the integration of equation (A) reduces to finding a *primitive* of its left-hand side.

Example. Let us find the general solution of the equation

$$(e^y + x)dx + (xe^y - 2y)dy = 0$$

As was shown (see pages 397 and 503), the left member of this equation is a total differential, namely, that of the function

$$u(x, y) = \frac{x^2}{2} + xe^y - y^2$$

Here we only need one particular primitive and therefore the arbitrary constant in the expression of $u(x, y)$ is omitted.

According to formula (*) the general solution of the given equation is expressed by the relation

$$\frac{x^2}{2} + xe^y - y^2 = C$$

representing y as an implicit function of x .

163. Approximate Methods of Solving First-order Differential Equations.

I. Direction field. If none of the particular methods of solving differential equations is applicable to a given equation of the first order

$$y' = f(x, y) \quad (*)$$

or if the integration involves too extensive calculations it is advisable to resort to some approximate methods of solution. Here we shall present Euler's graphical method and the numerical method of integration implied by it. Before proceeding to the study of these methods we shall discuss the geometric meaning of the first-order equation (*).

At every point $P(x, y)$ of the region in the xy -plane where the conditions of Cauchy's theorem (Sec. 158) hold equation (*) specifies the value of the slope of the

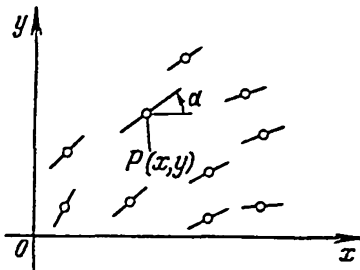


Fig. 231

tangent to the integral curve passing through the point $P(x, y)$. Thus, we have $\tan \alpha = y' = f(x, y)$ where α is the angle of inclination of the tangent to the axis Ox . This can be represented by a rectilinear arrow with slope equal to $f(x, y)$ passing through the point P (Fig. 231). Thus, equation (*) specifies in the plane Oxy the so-called *direction field*. For constructing

the direction field it is convenient to make use of the so-called *isoclines* of the field.

An *isocline* of the equation $y' = f(x, y)$ is a locus of points with the same constant slope ($y' = \text{const}$) of the field. The equation of an isocline is readily obtained if we substitute the corresponding value $y' = C$ into the differential equation:

$$C = f(x, y)$$

For arbitrary constant values of C this equation represents a *family of isoclines* of differential equation (*).

At all the points of an isocline corresponding to a given value of C the tangents to the integral curves have the same direction.

Consequently, the integration of a differential equation can be interpreted as the process of constructing the lines whose tangents at all the points go in the direction of the field.

As an example, consider the equation

$$y' = x^2 + y^2$$

This is a particular case of *Riccati's*¹ equation; it cannot be inte-

¹ J.F. Riccati (1676-1754), an Italian mathematician. Riccati's general equation is written as $y' = P(x)y^2 + Q(x)y + R(x)$. It was proved that, except some simple particular cases, its solution cannot be expressed in terms of integrals of elementary functions.

grated by means of any standard method given in the foregoing sections (generally, it is nonintegrable by quadratures).

To construct the direction field let us investigate the isoclines of the given equation. They obviously are the circles $x^2 + y^2 = C$ with centres at the origin. It is clear that we have $y' = 0$ only at the origin; at all the other points $y' > 0$. The value $y' = 0.5$ corresponds to the circle of radius ≈ 0.7 (for this isocline $\alpha \approx 26^\circ$). Similarly, on the circle of radius 1 we have $y' = 1$ (that is $\alpha = 45^\circ$), etc. The greater the radius of the circle, the greater the slope of the tangent (see Fig. 232).

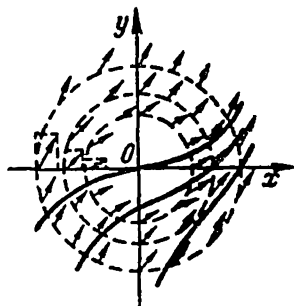


Fig. 232

In order to construct an integral curve we must take an initial point and draw a curve through it such that its tangent at each point goes in the direction of the field, that is, touches the corresponding arrow. A number of such curves are schematically shown in Fig. 232.

We now proceed to methods of *approximate integration*; in what follows we assume that the conditions of the existence and uniqueness theorem are fulfilled.

II. Euler's method. We shall consider Euler's method for constructing an integral curve of the first-order equation

$$y' = f(x, y) \quad (*)$$

passing through an initial point $M_0(x_0, y_0)$. This can be made approximately by means of simple constructions completely analogous to those used for the graphical integration of a function, that is for solving equation (*) in the particular case when $f(x, y)$ is a function solely dependent on x (see Sec. 95).

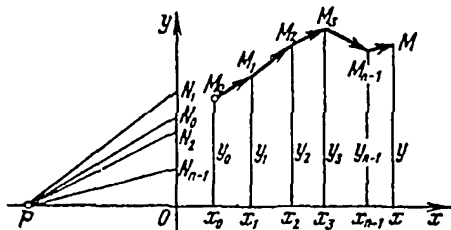


Fig. 233

Let us break up the interval $[x_0, x]$ into n parts with the aid of points $x_1, x_2, \dots, \dots, x_{n-1}$ (Fig. 233). Next,

through the points of division we draw straight lines parallel to Oy and then successively perform the following standard operations.

Compute the value of $f(x, y)$ at the point $M_0(x_0, y_0)$; according to equation (*), $f(x_0, y_0)$ is equal to the slope of the integral curve at the point $M_0(x_0, y_0)$. To construct the direction of the field at the point M_0 we take a point P on the x -axis to the left of the origin at a distance $OP = 1$ (the scale for measuring OP may

be different from that along the coordinate axes), lay off on the axis Oy a segment ON_0 whose length relative to the scale of OP is equal to the number $f(x_0, y_0)$ and join the point N_0 by a straight line to the point P . Then the direction of the segment PN_0 obviously coincides with that of the tangent to the integral curve at the point M_0 . Now draw from M_0 a line segment parallel to PN_0 to intersect the line $x=x_1$. We thus obtain a point M_1 which is taken as the point of the integral curve corresponding to $x=x_1$. This construction is equivalent to replacing the arc of the curve on the subinterval $[x_0, x_1]$ by a segment of its tangent at the initial point. Next we compute the value of $f(x, y)$ at the new point $M_1(x_1, y_1)$; this value $f(x_1, y_1)$ is equal to the slope of the tangent to the integral curve at the point M_1 . Setting, on the y -axis, a segment ON_1 of length equal, relative to OP , to the number $f(x_1, y_1)$ we join it with P by a straight line and then draw from the point M_1 a line segment parallel to PN_1 . This segment intersects the line $x=x_2$ at a point M_2 which we take as the point of the integral curve corresponding to $x=x_2$.

In the same manner we successively construct the points of the curve corresponding to the points of division $x_3, x_4, \dots$ of the interval $[x_0, x]$ until we arrive at the point $M(x, y)$. The polygonal line $M_0M_1M_2M_3 \dots M_{n-1}M$ thus obtained is taken as an approximate integral curve passing through the initial point $M_0(x_0, y_0)$.

III. Numerical integration. Now we describe analytically Euler's procedure for approximate integration of the differential equation (*).

The first operation yields the following relation between the coordinates of the points M_0 and M_1 :

$$y_1 - y_0 = f(x_0, y_0)(x_1 - x_0) \quad (1)$$

The second operation results in the analogous equality

$$y_2 - y_1 = f(x_1, y_1)(x_2 - x_1) \quad (2)$$

and so on; finally, the n th operation gives

$$y - y_{n-1} = f(x_{n-1}, y_{n-1})(x - x_{n-1}) \quad (n)$$

These n equalities make it possible to determine, in succession, the values of the unknown function at the points of division of the interval $[x_0, x]$. Indeed, the first equality specifies y_1 for the given x_0 and y_0 and the chosen x_1 , from the second equality, for the known x_1 and y_1 and the chosen x_2 we find the corresponding value y_2 , etc., until the desired value y is determined.

The smaller the maximum subinterval (that is, the greater the number n) and the closer the point x to x_0 , the more precise the ultimate result.

Thus, the practical computation of the approximate value of y can be carried out by means of equalities (1)-(n). This scheme is one of the *methods of numerical integration* of equation (*).

Example. Take the equation

$$y' = xy^2 + 1$$

It cannot be solved by any of the exact methods for integrating first-order equations. Let us find an approximate solution of this equation on the interval $[0, 1]$ satisfying the initial condition $y|_{x=0} = 0$ and compute y for $x = 1$.

Divide the interval $[0, 1]$ into four parts with the help of the points $x_0 = 0$, $x_1 = 0.25$, $x_2 = 0.5$, $x_3 = 0.75$ and $x_4 = 1$ and denote as y'_0 , y'_1 , y'_2 and y'_3 the values of y' at the points x_0 , x_1 , x_2 and x_3 , respectively. Since $x_0 = 0$ and $y_0 = 0$, we have $y'_0 = 1$, and hence

$$y_1 = 1(x_1 - x_0) = 0.25; \quad M_1(0.25, 0.25)$$

Furthermore,

$$y'_1 = 0.25 \cdot 0.25^2 + 1 = 1.016$$

Therefore

$$y_2 = 0.25 + 1.016 \cdot 0.25 = 0.504; \quad M_2(0.5, 0.504)$$

Then we find

$$y'_2 = 0.5 \cdot 0.504^2 + 1 = 1.127$$

$$y_3 = 0.504 + 1.127 \cdot 0.25 = 0.786 \quad \text{and} \quad M_3(0.75, 0.786)$$

Finally,

$$y'_3 = 0.75 \cdot 0.786^2 + 1 = 1.463$$

and

$$y = y_4 = 0.786 + 1.463 \cdot 0.25 = 1.152$$

Consequently, $y = 1.152$ is the sought-for approximation to the value at the point $x = 1$ of the particular solution of the given equation specified by the initial condition $y|_{x=0} = 0$.

There exist many more precise methods for approximate integration of differential equations; they are treated in special books (for instance, see [8], [12] and [13]).

164*. Singular Points and Singular Solutions. Envelope.

1. **Singular points.** Let there be given an equation $y' = f(x, y)$. Suppose that the function $f(x, y)$ satisfies the conditions of Cauchy's theorem on the existence and uniqueness of the solution (see Sec. 158) at all the points of a given region in the xy -plane except a point (x_0, y_0) . A point of this type is said to be a *singular point* of the differential equation. Through each point of this region, except the singular one, there is only one integral curve, and

our task is to investigate the behaviour of the integral curves passing through the points lying close to the singular point. Here we shall limit ourselves to demonstrating some typical examples without presenting the general investigation.

First of all, let us agree to consider the variables x and y as playing equivalent roles in the equation; this means that we can interpret y as a function of x or x as a function of y . This convention makes it necessary to state a more precise definition of a singular point. As an instance, take the equation $\frac{dy}{dx} = \frac{1}{y}$. The point $(0, 0)$ is its singular point since the right member of the equation is discontinuous for $y=0$. But if x is regarded as being dependent on y which is taken as an independent variable and if the equation is written in the form $\frac{dx}{dy} = y$ the point $(0, 0)$ is no longer singular because the right-hand side of the latter equation obviously satisfies all the conditions of Cauchy's theorem. For the initial point $(0, 0)$ it possesses the only solution $x = \frac{y^2}{2}$. The corresponding integral curve is a parabola tangent to the axis of ordinates at the origin. Thus, there is only one integral curve passing through the point $(0, 0)$, and therefore we have no reason to consider it as singular. The same refers to any other point of the axis of abscissas.

That is why we shall only consider as singular those points (x_0, y_0) at which the right members of both equations

$$\frac{dy}{dx} = f(x, y) \quad \text{and} \quad \frac{dx}{dy} = \frac{1}{f(x, y)}$$

are simultaneously discontinuous. For instance, this is the case for the equations

$$\frac{dy}{dx} = \frac{Ax + By}{Cx + Dy} \quad \text{and} \quad \frac{dx}{dy} = \frac{Cx + Dy}{Ax + By} \quad (*)$$

at the origin since both right members have no limits when x and y tend to zero in an arbitrary manner (see an analogous example on pp. 374, 375).

Below are several examples demonstrating the investigation of equations of type (*).

(1) $\frac{dy}{dx} = \frac{2y}{x}$. Separating the variables and integrating we obtain $\ln|y| = 2 \ln|x| + \ln|C|$ whence $y = Cx^2$. This is an equation of a family of parabolas. Now, writing the differential equation in the form $x dy - 2y dx = 0$ we find two more particular solutions: $y=0$ and $x=0$. In this example all the integral curves pass through

the singular point (i.e. through the origin; see Fig. 234). Such singular points are called *nodes* (*nodal points*).

Let the reader check that for the equation $\frac{dy}{dx} = \lambda \frac{y}{x}$ with $\lambda > 0$ the singular point $(0, 0)$ is also a node. In particular, for $\lambda = 1$ the family of integral curves consists of the straight lines passing through the origin.

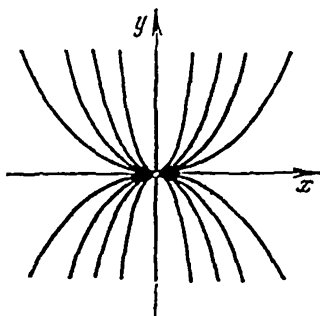


Fig. 234

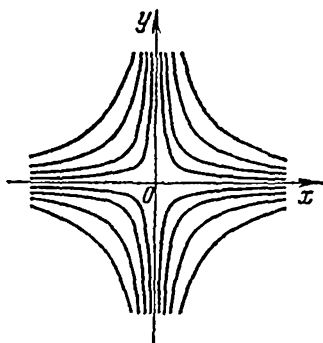


Fig. 235

(2) $\frac{dy}{dx} = -\frac{y}{x}$. The general solution of the equation is of the form $y = \frac{C}{x}$ and represents the family of the rectangular hyperbolas. To obtain all integral curves we must add to this family the axes of coordinates: $y=0$ and $x=0$ (Fig. 235). This is the so-called *saddle point* of the differential equation.

The situation is analogous for the equations $\frac{dy}{dx} = -\lambda \frac{y}{x}$ and $\frac{dy}{dx} = \lambda \frac{x}{y}$ when $\lambda > 0$.

(3) $\frac{dy}{dx} = -\frac{x}{y}$. The general solution is $x^2 + y^2 = C$ and the integral curves are the circles with centre at the origin (see Fig. 236). A singular point of this type is termed a *centre*; there is no integral curve passing through it. For the equation $\frac{dy}{dx} = -\lambda \frac{x}{y}$

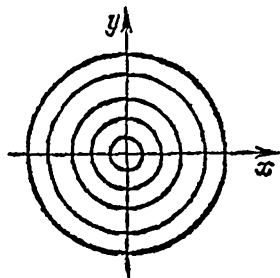


Fig. 236

where $\lambda > 0$ the singular point at the origin is also a centre; the family of integral curves is formed of the ellipses with centre at the origin.

(4) $\frac{dy}{dx} = \frac{x+y}{x-y}$. The substitution $y = ux$ brings the equation, after some simple transformations, to the equation with variables se-

parable $\frac{1-u}{1+u^2} du = \frac{dx}{x}$. On integrating and returning to y we obtain

$$\arctan \frac{y}{x} - \frac{1}{2} \ln \left(1 + \frac{y^2}{x^2} \right) + \ln C = \ln x$$

whence

$$\sqrt{x^2 + y^2} = Ce^{\arctan \frac{y}{x}}$$

In the polar coordinates with pole at the origin ($x = r \cos \varphi$, $y = r \sin \varphi$) the equation of the family takes the simpler form:

$$r = Ce^{\varphi}$$

This is the equation of logarithmic spirals (see Sec. 49) shown in Fig. 237. Here we have a singular point referred to as *focal*.

It can be proved (but we shall not go into particulars) that for equations (*) the origin is always a singular point of one of these four types, that is a *node*, a *saddle*, a *centre* or a *focus* provided that $AD - BC \neq 0$.

The investigation of singular points of more complicated equations is an extremely sophisticated problem which we shall not discuss in this course.

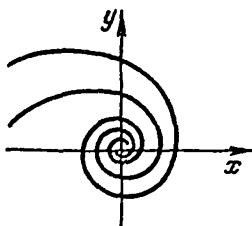


Fig. 237

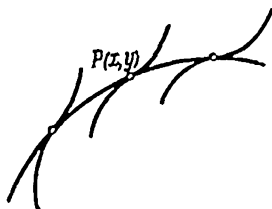


Fig. 238

II. Singular solutions. Envelope. A solution of the differential equation $y' = f(x, y)$ is said to be **singular** if at its every point the uniqueness of the solution of Cauchy's problem is violated. This means that in every neighbourhood of each point $P(x, y)$ of a singular solution there are at least two integral curves passing through the point P (see Fig. 238); at the point P these curves have one and the same tangent line whose slope is equal to $\tan \alpha = y' = f(x, y)$. The conditions of Cauchy's theorem are of course violated at all the points of a singular solution.

Let us begin with the simple equation $y' = 3y^{\frac{2}{3}}$. On separating the variables and integrating we arrive at the general solution $y = (x + C)^3$. Besides, the equation evidently has one more solution $y = 0$ which cannot be obtained from the general solution for

any value of C . It is clearly seen in Fig. 239 that all the integral curves (they are cubic parabolas) are tangent to the x -axis at the points of that axis. Hence, the function $y=0$ is a singular solution.

The appearance of this singular solution of the equation is accounted for by the fact that although the right member (the function $3y^{\frac{2}{3}}$) is continuous everywhere its derivative with respect

to y which is equal to $2y^{-\frac{1}{3}}$ for $y \neq 0$ does not exist for $y=0$ and hence one of the conditions of Cauchy's theorem is not fulfilled.

Now we proceed to the general case. Let a differential equation $y' = f(x, y)$ have a general solution $y = \varphi(x, C)$ (or a general integral $\Phi(x, y, C) = 0$) describing a family of integral curves. It may happen that there exists a curve tangent at its every point to an integral curve of the family, its every arc being tangent to an infinite number of curves of the family. A curve of this kind is called an *envelope* of the family of integral curves. The envelope (provided it exists) is also an integral curve of the given equation since at every point it is tangent to an integral curve of the family and hence its tangent goes in the direction of the corresponding field (see Sec. 163, I). At each point of the envelope the uniqueness of the solution of the Cauchy problem is violated.

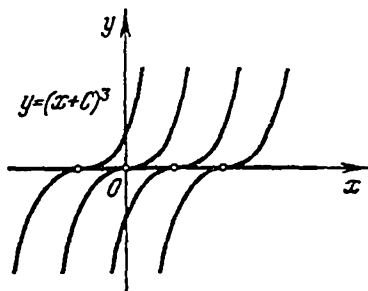


Fig. 239

We deal with envelopes of families of curves not only when investigating singular solutions of differential equations but also in some other problems. For instance, in Sec. 67 we showed that the evolute of a plane curve is the envelope of the family of normals to that curve (which is the evolute).

III. Envelope of a family of curves. Let $F(x, y, C) = 0$ be an equation describing a family of curves dependent on one parameter (a *one-parameter family of curves*). We shall suppose that the function $F(x, y, C)$ possesses continuous partial derivatives with respect to all its arguments and that none of the curves of the family has singular points¹. Let the given family have an envelope tan-

¹ For this condition to hold it is sufficient that at every point of any curve $F(x, y, C) = 0$ the partial derivatives $F'_x(x, y, C)$ and $F'_y(x, y, C)$ do not vanish simultaneously.

gent to every curve of the family at exactly one point (Fig. 240). The coordinates x and y of such a point P are completely determined by the curve of the family to which that point belongs, and the curve itself, in its turn, is specified by the corresponding value of the parameter C . Therefore we can write

$$x = x(C), \quad y = y(C) \quad (*)$$

When the parameter C is varied the point P moves along the envelope; consequently equations $(*)$ give a parametric representation

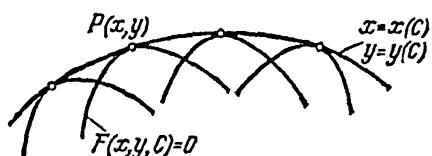


Fig. 240

of the envelope. It is of course supposed that the envelope is a smooth curve, which means that the functions $(*)$ are continuous and possess continuous derivatives not turning to zero simultaneously.

Let us write down the condition expressing the fact that the envelope and the corresponding curve of the family have a common tangent at the point P . The slope of the tangent to the envelope is obtained by the rule for differentiating a function represented parametrically: $k_{env} = y'_x = \frac{y'_C}{x'_C}$. The slope of the tangent to the curve of the family is found according to the rule of differentiation of an implicit function applied to the function specified by the equation $F(x, y, C) = 0$ for the *constant* value of the parameter C : $k_{curve} = \frac{dy}{dx} = -\frac{F'_x}{F'_y}$.

Equating these slopes we obtain

$$\frac{y'_C}{x'_C} = -\frac{F'_x}{F'_y}, \quad \text{i.e.} \quad F'_x x'_C + F'_y y'_C = 0$$

Furthermore, for any value of the parameter C the point of the envelope having the coordinates $x(C), y(C)$ belongs to the curve of the family specified by the same value of C , and therefore the equality

$$F[x(C), y(C), C] = 0$$

must be fulfilled for all C . The total derivative with respect to C of the left-hand side of the equality is thus equal to zero:

$$F'_x x'_C + F'_y y'_C + F'_C = 0$$

Now taking into account the equality of the slopes we obtain $F'_C = 0$.

Thus, if the envelope exists the functions (*) representing it parametrically must satisfy the relations

$$F(x, y, C) = 0, F'_C(x, y, C) = 0 \quad (**)$$

Since it is unknown in advance whether there exists an envelope the following assertion which we give without proof is of great importance:

*If the function $F(x, y, C)$ satisfies the above conditions and if it is possible to find from equations (**) functions $x(C)$ and $y(C)$ which yield parametric equations of a smooth curve this curve is the envelope of the family of curves $F(x, y, C) = 0$.*

If the family of curves is given by an equation $y = \varphi(x, C)$ the equations (**) are simplified:

$$y = \varphi(x, C), \varphi'_C(x, C) = 0 \quad (***)$$

Examples. (1) Let us find the envelope of the family of circles

$$(x - C)^2 + y^2 = 1$$

The differentiation with respect to C yields $2(x - C) = 0$. Equations (**) then reduce to one equation $y^2 = 1$ whence $y = 1$ or $y = -1$. Consequently, the envelope consists of two straight lines parallel to Ox (see Fig. 241).

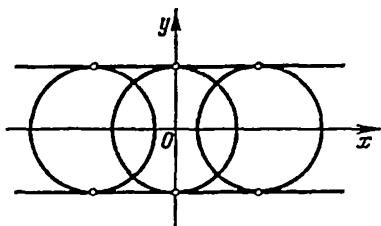


Fig. 241

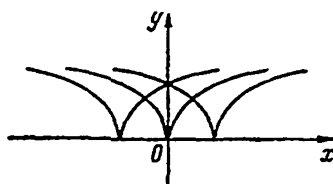


Fig. 242

(2) Consider the envelope of the family $y = (x + C)^2$; this is the family of integral curves of the differential equation considered at the beginning of Sec. II. Here equations (**) reduce to the equation $y = 0$; hence, the envelope is the x -axis (see Fig. 239).

In this example the envelope intersects all the curves of the family and is tangent to them at the points of intersection.

If the function $F(x, y, C)$ does not satisfy the above conditions equations (**) may not represent an envelope. For instance, take the family of semicubical parabolas $y^3 - (x + C)^2 = 0$ (Fig. 242). Here equations (**) reduce to one equation $y = 0$. It is clearly seen that the axis Ox is not an envelope but a line consisting of the *cusps* of the curves of the family and thus is its *cusp locus*.

In conclusion let us consider *Clairaut's*¹ differential equation

$$y = xy' + \varphi(y')$$

Its general solution is readily found:

$$y = Cx + \varphi(C)$$

Indeed, this formula implies $y' = C$, and the substitution into the equation leads to the identity $Cx + \varphi(C) = x(C) + \varphi(C)$. Geometrically, this general solution is a family of straight lines.

In problems leading to Clairaut's equation we most often are interested in its singular solution which is the envelope of the family of straight lines. This solution is given by the parametric equations $y = Cx + \varphi(C)$, $x + \varphi'(C) = 0$, that is

$$x = -\varphi'(C), \quad y = -C\varphi'(C) + \varphi(C)$$

Example. The equation $y = xy' - \frac{y'^2}{4}$ has the general solution $y = Cx - \frac{C^2}{4}$. Eliminating the parameter C from this equation and from the equation $x - \frac{C}{2} = 0$ we obtain the explicit equation of the envelope: $y = x^2$. This is the sought-for singular solution. Its graph is a parabola, and the general solution is the family of tangent lines to the parabola.

§ 2. Differential Equations of the Second and Higher Orders

165. Differential Equations of the Second Order. We now proceed to studying the differential equations of the form

$$F(x, y, y', y'') = 0$$

whose left-hand side involves the *second derivative* of the unknown function. They are called *differential equations of the second order*.

Usually we deal with equations *resolved* with respect to the second derivative:

$$y'' = f(x, y, y') \quad (*)$$

We shall begin with a simple example:

$$y'' = x$$

Integrating twice we find in succession the first derivative $y' = \frac{x^2}{2} + C_1$ and the function y itself:

$$y = \int \left(\frac{x^2}{2} + C_1 \right) dx + C_2 = \frac{x^3}{6} + C_1 x + C_2$$

¹ A.C. Clairaut (1713-1765), a French mathematician.

Since we have integrated twice the resultant expression contains two arbitrary constants which are denoted by C_1 and C_2 .

Now we state without proof the following general conclusion: *the second-order differential equation $y''=f(x, y, y')$ usually possesses an infinitude of solutions determined by a formula $y=\varphi(x, C_1, C_2)$ containing two arbitrary constants.* Such a family of solutions is called the *general solution of the equation*.

To isolate a particular solution we set initial conditions which usually are of the form

$$y|_{x=x_0}=y_0, \quad y'|_{x=x_0}=y'_0$$

where y_0 and y'_0 are some given numbers.

For example, let us determine the particular solution of the above equation $y''=x$ satisfying the initial conditions $y|_{x=2}=2$, $y'|_{x=2}=3$ (Cauchy's problem). Substituting the given values into the general solution and into the expression of its derivative we obtain the system of equations

$$3=2+C_1, \quad 2=\frac{4}{3}+2C_1+C_2$$

whence

$$C_1=1 \quad \text{and} \quad C_2=-\frac{4}{3}$$

Consequently, the desired particular solution is

$$y=\frac{x^3}{6}+x-\frac{4}{3}$$

The geometrical meaning of the initial conditions is that besides a point (x_0, y_0) through which the sought-for integral curve must pass we also set the slope y'_0 of the tangent to the curve at that point. It should be stressed that, since the general solution of a second-order equation depends on two arbitrary constants, *there are infinitely many integral curves passing through the given point*, and only one of them has the *preassigned* slope. For instance,

it is readily seen that if we put $C_2=\frac{2}{3}-2C_1$ in the general solution of the above example the graph of the function $y=\frac{x^3}{6}+C_1x+\left(\frac{2}{3}-2C_1\right)$ thus obtained passes through the point $(2, 2)$ for any value of C_1 . Selecting from all these curves the one with the slope 3 at that point we arrive at the desired particular solution.

To formulate the existence theorem for equations of the second order (*) we shall consider the right member of the equation as a function of the three independent variables x, y and y' because when setting the initial conditions we independently choose the

coordinates x_0, y_0 and the slope of the tangent y'_0 which are not related to each other by any condition.

Theorem (On the Existence and Uniqueness of the Solution of the Differential Equation of the Second Order). If the function $f(x, y, y')$ and its partial derivatives $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial y'}$ are continuous in some neighbourhood of the point (x_0, y_0, y'_0) the equation $y'' = f(x, y, y')$ possesses a unique solution $y = y(x)$ defined in a sufficiently small interval $[x_0 - h, x_0 + h]$ and satisfying the initial conditions $y(x_0) = y_0$ and $y'(x_0) = y'_0$.

This theorem shows immediately that the equation $y'' = \frac{y'}{x} + y$ has a unique solution for the initial conditions $y|_{x=1} = 2, y'|_{x=1} = -1$. Of course, the question of how this solution can be found remains open. If for the same equation initial conditions are set at the point $x=0$ the existence theorem does not enable us to draw any conclusion since for these initial values the right member of the equation is not defined.

As in the case of a first-order equation, the problem of determining the particular solution of a differential equation of the second order for given initial conditions is referred to as *Cauchy's problem*.

In the case of a second-order equation a particular solution can also be isolated by setting the so-called *boundary conditions* which usually specify the values of the function y at two different points:

$$y|_{x=x_1} = y_1 \quad \text{and} \quad y|_{x=x_2} = y_2$$

A problem of this kind can be encountered in the course of strength of materials in connection with studying the deflection of a beam. More general *boundary-value problems* are dealt with in mathematical physics.

As an example, let us investigate the solution of the equation $y'' = x$ satisfying the boundary conditions $y|_{x=1} = 0$ and $y|_{x=2} = 0$. As was shown, the general solution of the equation is $y = \frac{x^3}{6} + C_1x + C_2$. Substituting the given values into this expression we obtain a system of two equations for determining the arbitrary constants:

$$\frac{1}{6} + C_1 + C_2 = 0, \quad \frac{8}{3} + 2C_1 + C_2 = 0$$

From this system we find $C_1 = -\frac{7}{6}$ and $C_2 = 1$ whence the desired particular solution is $y = \frac{x^3}{6} - \frac{7}{6}x + 1$.

In this example we have found the unique particular solution satisfying the given boundary conditions. But this is not always the case; it may turn out that a second-order equation has no

solution satisfying given boundary conditions and in other cases it may possess an infinitude of such solutions. An example illustrating what has been said will be given at the end of Sec. 170. It is these properties that characterize an essential distinction between boundary-value problems and initial-value problems: if some initial values are given the existence theorem (provided its conditions are fulfilled) immediately guarantees the uniqueness of the solution of the problem while for given boundary conditions we can only judge upon the existence and the uniqueness of the solution after the general solution of the equation has been found.

166. Some Particular Types of Equations of the Second Order. Let us consider some particular types of the second-order equation

$$y'' = f(x, y, y')$$

which can readily be reduced to equations of the first order.

I. The right-hand side of the equation does not contain y and y' :

$$y'' = f(x) \quad (*)$$

Since $y'' = (y')'$, we have

$$y' = \int f(x) dx + C_1$$

On integrating once again we obtain

$$y = \int \left[\int f(x) dx \right] dx + C_1 x + C_2$$

where C_1 and C_2 are arbitrary constants. An example of such an equation was considered in Sec. 165.

II. The right-hand side of the equation does not contain y :

$$y'' = f(x, y') \quad (**)$$

Let us put $y' = z$; then $y'' = z'$, and equation (**) becomes a first-order equation with respect to z :

$$z' = f(x, z)$$

If the solution $z = \varphi(x, C_1)$ of this equation is found the sought-for solution of the original equation is obtained by the integration of the equality $y' = z$, that is

$$y = \int \varphi(x, C_1) dx + C_2$$

Example. Solve the equation

$$y'' + \frac{y'}{x} = x$$

Putting $y' = z$ and $y'' = z'$ we arrive at the equation of the first order $z' + \frac{z}{x} = x$ which is linear. On solving this equation we find $z = \frac{x^2}{3} + \frac{C_1}{x}$. Finally, $y' = \frac{x^2}{3} + \frac{C_1}{x}$ and $y = \frac{x^3}{9} + C_1 \ln|x| + C_2$.

III. The right-hand side of the equation does not contain

$$y'' = f(y, y') \quad (*)$$

Let us put $y' = p$ and regard p as a function of y . The differentiation of this relation results in $y'' = \frac{dp}{dx}$. In order to eliminate x we use the transformation

$$\frac{dp}{dx} = \frac{dp}{dy} \frac{dy}{dx} = \frac{dp}{dy} y' = \frac{dp}{dy} p$$

Thus we obtain

$$y'' = p \frac{dp}{dy}$$

which leads to

$$p \frac{dp}{dy} = f(y, p)$$

that is to an equation of the first order relative to p as function of y . If the solution $p = \varphi(y, C_1)$ of this equation is determined, the desired solution of the original equation is found from equation with variables separable:

$$\frac{dy}{dx} = p = \varphi(y, C_1), \quad \text{that is } \frac{dy}{\varphi(y, C_1)} = dx$$

and

$$\int \frac{dy}{\varphi(y, C_1)} = x + C_2$$

Example. Let us solve the equation

$$2yy'' + y'^2 = 0$$

Putting $y' = p$ and $y'' = p \frac{dp}{dy}$ we obtain

$$2yp \frac{dp}{dy} + p^2 = 0$$

This is a first-order equation with variables separable. On bringing it to the form $\frac{dp}{p} = -\frac{dy}{2y}$ and integrating we get $\ln |p| = -\frac{1}{2} \ln |y| + \ln |C_1|$ and $p = \frac{C_1}{\sqrt{y}}$ where $y > 0$. Now we find from the equation $\frac{dy}{dx} = \frac{C_1}{\sqrt{y}}$ and obtain the required solution $y^{\frac{3}{2}} = C_1 x + C_2$ or $y = (C_1 x + C_2)^{\frac{2}{3}}$. When cancelling by p we obtain the solution $p = y' = 0$ of the transformed equation, that is the solution $y = \text{const}$ of the given equation.

In both cases II and III we introduced y' as a new auxiliary function and obtained in this way an equation of the first order

If an equation of the second order is of the form $y'' = f(y')$ to which both methods II and III are applicable one should follow the way which is more convenient for the given particular case.

167. Applications to Mechanics.

1. **Catenary.** Let us consider a uniform nonextensible flexible thread suspended by its ends (see Fig. 243). We shall investigate the shape of the thread under the action of its own weight.

Let us draw the axis Oy vertically through the lowest point N of the line and the axis Ox horizontally at a certain distance (which will be specified later on from the point N).

Take an arbitrary point M of the line. In the state of equilibrium the portion NM of the line can be regarded as a rigid body subjected to the action of three forces, namely, the force of horizontal tension H , the force of tension T applied at the point M and directed along the tangent to the line and the force of its own weight P equal to $s\delta$ where s is the length of the arc NM and δ the weight of the thread per unit length.

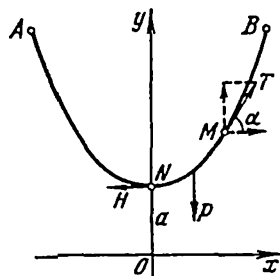


Fig. 243

The equilibrium conditions imply

$$T \sin \alpha = s\delta \quad \text{and} \quad T \cos \alpha = H$$

Dividing the first equality by the second we obtain

$$\tan \alpha = \frac{\delta}{H} s$$

Consequently, if $y = \varphi(x)$ is the sought-for equation of the curve ANB we must have

$$y' = ks \quad \text{where} \quad k = \frac{\delta}{H} = \text{const}$$

Let us differentiate this equality with respect to x :

$$y'' = ks' = k\sqrt{1+y'^2}$$

This equation can be solved as was indicated in Sec. 166, II. We put $y' = z$; then $y'' = z'$, and we thus obtain

$$z' = k\sqrt{1+z^2}, \quad \text{that is} \quad \frac{dz}{\sqrt{1+z^2}} = k dx$$

It follows that

$$\ln(z + \sqrt{1+z^2}) = kx + C_1$$

At the point N we have $x=0$ and $z=y'=0$ (since N is the lowest point of the line at which the tangent is horizontal). Con-

sequently, $C_1 = 0$ and $z + \sqrt{1 + z^2} = e^{kx}$. Transposing z to the right side and squaring we arrive at the relation

$$z = y' = \frac{1}{2}(e^{kx} - e^{-kx}) = \sinh kx$$

The integration of this relation gives

$$y + C_2 = \frac{1}{2k}(e^{kx} + e^{-kx}) = \frac{1}{k} \cosh kx$$

Let us now choose the distance ON equal to $\frac{1}{k}$. Then $C_2 = 0$, and we derive the desired equation of the line:

$$y = \frac{1}{2k}(e^{kx} + e^{-kx}) = \frac{1}{k} \cosh kx$$

This example illustrates how we can determine in succession the values of the constants of integration. If we tried to begin with finding the general solution this would involve more extensive calculations.

Introducing the notation $\frac{1}{k} = a$ we write

$$y = \frac{a}{2} \left(e^{\frac{x}{a}} + e^{-\frac{x}{a}} \right) = a \cosh \frac{x}{a}$$

This is the equation of a *catenary* whose shape is taken by the freely suspended thread.

II. Some problems of particle dynamics. Let a material point be in a rectilinear motion under the action of a force directed along the trajectory. By Newton's second law, the product of the mass m by the acceleration w is equal to the acting force F : $mw = F$.¹ Since $w = \frac{d^2s}{dt^2}$ where t is time and s the path length we obtain the differential equation of the second order for the law of motion:

$$m \frac{d^2s}{dt^2} = F$$

In a mechanical problem of a general type the force F may depend on time t and also on the path length s and the velocity v , i.e.

¹ Here we consider the acceleration w and the force F not as vectors but as scalars since the conditions of the problem imply that the directions of the force and of the acceleration are solely determined by their algebraic signs.

In the general case the equation of motion is written in the form $m \frac{d^2\mathbf{r}}{dt^2} = \mathbf{F}$ where $\mathbf{r}(t)$ is the radius-vector of the moving point (see Sec. 70). Projecting this equality on the coordinate axes we obtain a system of three differential equations whose integration encounters essential difficulties.

$F = F\left(t, s, \frac{ds}{dt}\right)$. In mechanics the differentiation with respect to time is often marked not by prime but by the dot: $v = \dot{s}$ and $w = \ddot{s}$. Then the differential equation of motion takes the form

$$m\ddot{s} = F(t, s, \dot{s}) \quad (*)$$

The initial conditions are usually written as $s|_{t=0} = s_0$ and $\dot{s}|_{t=0} = \dot{s}_0$, their mechanical meaning being quite obvious: we set the initial position and the initial velocity of the point. Let us consider some examples.

(1) **Uniformly accelerated motion.** Let the force F be constant. Denote the ratio $\frac{F}{m}$ as a , then $\ddot{s} = a$. Integrating we obtain $\dot{s} = at + C_1$ and $s = \frac{at^2}{2} + C_1t + C_2$. It is obvious that $C_1 = \dot{s}|_{t=0} = v_0$ and $C_2 = s|_{t=0} = s_0$. Thus, we have derived the well-known formula for the distance travelled in a uniformly accelerated motion: $s = \frac{at^2}{2} + v_0t + s_0$.

(2) **Motion of a point in a medium with resistance.** Experiments show that every body moving in a medium undergoes the resistance of the medium. The force of resistance increases together with the velocity and depends both on the properties of the medium (i.e. on its density, viscosity, etc.) and on the dimensions and the shape of the body. If the velocity of motion is not high and the dimensions of the body are not large the force of resistance can be approximately regarded as being directly proportional to the velocity v :

$$F_{res} = -kv$$

Here $k > 0$, and the minus sign in front of this proportionality factor indicates that the direction of the force of resistance of the medium is always opposite to the motion¹. When the velocity of motion is high the force of resistance becomes proportional to the square of the velocity:

$$F_{res} = -\gamma v^2$$

In this case the proportionality coefficient γ is computed by the formula

$$\gamma = C\rho S$$

where ρ is the density of the medium, S is the *maximum cross-section (mid-section)* perpendicular to the direction of motion and C

¹ For instance, according to Stokes' law, for the motion of a ball of radius R in a fluid, $k = 6\pi R\eta$ where η is the viscosity of the medium.

is a dimensionless factor determined experimentally and dependent on the geometric form of the moving body¹.

The question of what kind of law of resistance should be chosen in a concrete situation is of course decided in experiments.

As an example, let us consider a body falling on the earth and acted upon by gravitation and air drag.

Let the mass of the body be m and its initial velocity zero. The body undergoes the action of the force of gravity mg and of the force of resistance which is usually assumed to be proportional to the square of the velocity of motion. In this case the differential equation of motion takes the form

$$m\ddot{s} = mg - \gamma \dot{s}^2$$

Making the substitution

$$\dot{s} = v, \quad \ddot{s} = \frac{dv}{dt}$$

we arrive at the first-order equation

$$m \frac{dv}{dt} = mg - \gamma v^2$$

Rewriting this equation as

$$\frac{dv}{dt} = \frac{\gamma}{m} \left(\frac{mg}{\gamma} - v^2 \right)$$

putting $\frac{\gamma}{m} = a$, $\frac{mg}{\gamma} = b^2$ and separating the variables we obtain

$$\frac{dv}{b^2 - v^2} = a dt$$

On integrating (see Formula 11 in the table of integrals), we receive

$$\frac{1}{2b} \ln \frac{b+v}{b-v} = at + C$$

Since $v|_{t=0} = 0$, we have $C = 0$. Then $\frac{b+v}{b-v} = e^{2abt}$ and

$$v = b \frac{e^{2abt} - 1}{e^{2abt} + 1} = b \tanh(abt)$$

The formula obtained shows that the velocity v is always less than b and tends to this value as $t \rightarrow \infty$. Hence, the velocity does not increase indefinitely (which is the case in a free fall in vacuum) and tends to a definite limit referred to as the *terminal velocity of fall*:

$$v_{term} = b = \sqrt{\frac{mg}{\gamma}}$$

¹ The coefficient C is usually determined with the aid of a wind tunnel in which an air flow is incident on the motionless body.

Practically the terminal velocity is achieved comparatively quickly; the value of the terminal velocity and also the time taken to achieve it increase together with the mass of the body and with the reduction of the resistance coefficient γ , which is coherent with our everyday experience and intuition.

Now we shall establish the relationship between the velocity of fall and the distance travelled. As is known, for the free fall (when air resistance is neglected) we have $v = \sqrt{2gs}$ where s is the distance passed in the fall. To discover the sought-for relationship we can proceed from the known law of variation of the velocity to find the distance travelled by means of integration and then determine the connection between v and s . However, we shall follow a simpler way, namely, eliminate the independent variable t from the differential equation of motion. To this end we use the substitution indicated in Sec. 166, III:

$$\dot{s} = v, \quad \ddot{s} = \frac{dv}{dt} = \frac{dv}{ds} \frac{ds}{dt} = v \frac{dv}{ds}$$

Now, making use of the new notation we write down the differential equation in the form

$$v \frac{dv}{ds} = a(b^2 - v^2)$$

On separating the variables and integrating, we receive

$$-\frac{1}{2} \ln(b^2 - v^2) = as + C$$

Since $v|_{s=0} = 0$, we have $C = -\frac{1}{2} \ln b^2$, and, consequently, $as = \frac{1}{2} \ln \frac{b^2}{b^2 - v^2}$ whence $v = b \sqrt{1 - e^{-2as}}$. Replacing a and b by their values we obtain

$$v = \sqrt{\frac{mg}{\gamma}} \sqrt{1 - e^{-\frac{2\gamma s}{m}}}$$

It follows from this formula that if the ratio $\frac{\gamma}{m}$ is large. (this means that we deal with a bluff body, not streamlined), the quantity $e^{-\frac{2\gamma s}{m}}$ can be neglected even for not too large values of s . This approximation again leads to the expression $\sqrt{\frac{mg}{\gamma}}$ for the terminal velocity. If the ratio $\frac{\gamma}{m}$ is very small we can use the approximate formula $e^x \approx 1 + x$ (applicable for small x ; see Sec. 36) to obtain, for small s , the relation

$$v \approx \sqrt{\frac{mg}{\gamma}} \sqrt{1 - \left(1 - \frac{2\gamma s}{m}\right)} = \sqrt{2gs}$$

which means that in these circumstances the formula for the velocity of free fall can be used as an approximation.

Let the reader analyse the free fall of a body in a medium with force of resistance directly proportional to the velocity (this, for instance, is the case when the body is falling in water) and show that for this motion as well the velocity v has a definite terminal (limiting) value as t increases.

Later on, in Sec. 175, we shall come back to problems of dynamics and consider harmonic oscillations of a point.

168. Higher-order Differential Equations. The problem of solving a differential equation becomes more complicated as the left-hand side of the equation involves derivatives of higher order. In this section we shall limit ourselves to stating some basic definitions related to differential equations of the n th order.

Definition. The *order* of a differential equation is the highest order of the derivative involved in the equation.

Most often we deal with equations resolved with respect to the highest derivative. The general form of a differential equation of the n th order solved for the n th derivative is

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}) \quad (*)$$

The general solution of such an equation depends on n arbitrary constants: $y = \varphi(x, C_1, C_2, \dots, C_n)$. To isolate a particular solution corresponding to a concrete problem one must set *initial conditions*; in *Cauchy's problem* they are of the form

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y'_0, \quad \dots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}$$

that is for $x = x_0$ certain values of the function y itself and of its derivatives up to order $n-1$ inclusive are assigned. Differentiating the general solution $n-1$ times and substituting the initial data we obtain a system of n equations with the n unknowns $C_1, C_2, \dots, C_n$.

As to the existence and uniqueness of a particular solution, this question is answered by the theorem below which is analogous to the theorems stated previously for the cases $n=1$ and $n=2$. In the formulation of this theorem the right-hand side of equation (*) (the function f) is regarded as a function of $n+1$ independent variables $x, y, y', \dots, y^{(n-1)}$; if the function f is continuous and possesses continuous partial derivatives with respect to the variables $y, y', \dots, y^{(n-1)}$ in a neighbourhood of the point $(x_0, y_0, y'_0, \dots, y_0^{(n-1)})$, equation (*) has a unique solution satisfying the initial conditions $y|_{x=x_0} = y_0, y'|_{x=x_0} = y'_0, \dots, y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}$ and defined in a sufficiently small interval $[x_0 - h, x_0 + h]$.

Consider a simple type of n th-order equation which is easily solved for any n :

$$y^{(n)} = f(x)$$

The general solution of this equation is found by means of n successive integrations; every integration results in the appearance of an arbitrary constant.

Example. Let us solve the equation $y'''' = \sin x$. We obtain in succession

$$\begin{aligned}y''' &= -\cos x + C_1 \\y'' &= -\sin x + C_1 x + C_2 \\y' &= \cos x + C_1 \frac{x^2}{2} + C_2 x + C_3\end{aligned}$$

and, finally,

$$y = \sin x + C_1 \frac{x^3}{6} + C_2 \frac{x^2}{2} + C_3 x + C_4$$

Changing the notation of the arbitrary constants we can rewrite the general solution in the form

$$y = \sin x + C_1 x^3 + C_2 x^2 + C_3 x + C_4$$

§ 3. Linear Differential Equations

169. Linear Equations of the Second Order. General Properties.

Definition. A *linear differential equation* of the second order is an equation which involves linearly the unknown function and its derivatives or, equivalently, which is an equation of the first degree with respect to the unknown function and its derivatives.

The general form of such an equation solved with respect to the highest derivative is

$$y'' + a_1(x)y' + a_2(x)y = f(x) \quad (*)$$

The right member of this equation (the function $f(x)$) is referred to as the *nonhomogeneity term*. If the function $f(x)$ is identically equal to zero, equation (*) is called a *homogeneous linear equation*; otherwise it is called a *nonhomogeneous linear equation*.

If the functions $a_1(x)$, $a_2(x)$ and $f(x)$ are continuous in an interval $a < x < b$ (finite or infinite), equation (*) possesses a unique solution satisfying arbitrary initial conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y'_0 \quad \text{where } x_0 \in (a, b)$$

This is implied by the fact that equation (*) written in the form $y'' = -a_1(x)y' - a_2(x)y + f(x)$ satisfies all the conditions of the existence and uniqueness theorem stated in Sec. 165 since the right-hand side of the latter equation and its partial derivatives with respect to y and y' (which are, respectively, equal to $-a_2(x)$ and $-a_1(x)$) are continuous. Moreover, it can be proved that the solution of the linear equation determined by the initial conditions

exists and is unique on the whole interval (a, b) (remember that in the general case of an arbitrary nonlinear equation the theorem only guarantees the existence and uniqueness of the solution for a sufficiently small interval $[x_0 - h, x_0 + h]$).

Note that the linear equation

$$p_0(x)y'' + p_1(x)y' + p_2(x)y = f(x)$$

can be reduced to form (*) by dividing both sides by $p_0(x)$. But then at the points where $p_0(x) = 0$ the conditions of the existence theorem can be violated; such points are called *singular*¹.

I. **Homogeneous Linear Equations.** We shall begin with a homogeneous linear equation

$$y'' + a_1(x)y' + a_2(x)y = 0 \quad (**)$$

It is supposed that the functions $a_1(x)$ and $a_2(x)$ are continuous in an interval (a, b) which can be finite or infinite. For brevity, we shall sometimes write a_1 and a_2 instead of $a_1(x)$ and $a_2(x)$; in particular cases a_1 and a_2 can be constant.

Every homogeneous linear equation always possesses the so-called *trivial (zero)* solution $y \equiv 0$.

Let us discuss some properties of solutions of equation (**).

Theorem I. If $y_1(x)$ and $y_2(x)$ are two solutions of homogeneous linear equation (**) the function $y = C_1y_1(x) + C_2y_2(x)$ is also a solution of the equation for any constants C_1 and C_2 .

For brevity, we shall write these solutions as y_1 and y_2 ; the expression $C_1y_1 + C_2y_2$ is called their *linear combination*.

Proof. Differentiate the function $y = C_1y_1 + C_2y_2$ twice:

$$y' = C_1y_1' + C_2y_2', \quad y'' = C_1y_1'' + C_2y_2''$$

Now, substituting y , y' and y'' into the left-hand side of equation (**) we obtain

$$\begin{aligned} C_1y_1'' + C_2y_2'' + a_1(C_1y_1' + C_2y_2') + a_2(C_1y_1 + C_2y_2) = \\ = C_1(y_1'' + a_1y_1' + a_2y_1) + C_2(y_2'' + a_1y_2' + a_2y_2) \end{aligned}$$

The expressions in the parentheses are the results of the substitution of the functions y_1 and y_2 in equation (**); since these functions are solutions of (**) both expressions are identically equal to zero, and hence the function $y = C_1y_1 + C_2y_2$ does in fact satisfy equation (**). The theorem has been proved.

The next theorem indicates the conditions under which a linear combination of two solutions is the general solution.

Theorem II. If y_1 and y_2 are solutions of equation (**) such that their ratio is not equal to a constant $\left(\frac{y_2}{y_1} \neq \text{const}\right)$ the linear

¹ The investigation of singular points of some special types of linear equation plays an important role in mathematical physics.

combination of these functions

$$y = C_1 y_1 + C_2 y_2 \quad (***)$$

is the general solution in the sense that every solution of equation (**) is expressible in the form (***) for appropriately chosen constants C_1 and C_2 .

The condition $\frac{y_2}{y_1} \neq \text{const}$ indicates that neither of these solutions is a trivial one.

Proof. According to Theorem I, the function y is a solution for any C_1 and C_2 . Let us prove that under the above hypothesis concerning y_1 and y_2 there always exist, for any initial conditions $y|_{x=x_0} = y_0$, $y'|_{x=x_0} = y'_0$, the constants C_1 and C_2 such that the solution expressed by the linear combination of y_1 and y_2 with the coefficients C_1 and C_2 satisfies these initial conditions. The point x_0 must of course belong to the interval on which the equation is considered; in particular, if a_1 and a_2 are constant x_0 can be any point.

To prove this assertion let us substitute the initial values into the expressions of the function y and its derivative. This results in the following system of linear algebraic equations with respect to C_1 and C_2 :

$$\begin{aligned} C_1 y_{10} + C_2 y_{20} &= y_0 \\ C_1 y'_{10} + C_2 y'_{20} &= y'_0 \end{aligned}$$

where the values $y_{10} = y_1(x_0)$, $y'_{10} = y'_1(x_0)$, $y_{20} = y_2(x_0)$ and $y'_{20} = y'_2(x_0)$ are known and C_1 and C_2 are to be found.

Now we shall show that the determinant of this system is non-zero:

$$W_0 = \begin{vmatrix} y_{10} & y_{20} \\ y'_{10} & y'_{20} \end{vmatrix} \neq 0$$

This will imply that the system possesses a unique solution for any right-hand terms.

Assume the contrary: let the determinant be equal to zero. Then the homogeneous linear system of algebraic equations

$$C_1 y_{10} + C_2 y_{20} = 0, \quad C_1 y'_{10} + C_2 y'_{20} = 0$$

which corresponds to the zero initial conditions $y_0 = 0$, $y'_0 = 0$ possesses, besides the zero solution $C_1 = 0$, $C_2 = 0$, an infinite number of nonzero solutions. If C_{10} , C_{20} is one of them, the function $C_{10} y_1 + C_{20} y_2$ is a solution of equation (**) corresponding to the zero initial conditions. But since the trivial solution $y \equiv 0$ obviously satisfies the same conditions, we conclude, by the uniqueness theorem, that

$$C_{10} y_1 + C_{20} y_2 \equiv 0$$

because there can only be one such solution. It now follows that $\frac{y_2}{y_1} = -\frac{C_{10}}{C_{20}} = \text{const}$, which contradicts the condition $\frac{y_2}{y_1} \neq \text{const}$ (if $C_{20} = 0$ then $C_{10} \neq 0$ and $y_1 \equiv 0$, which, according to the hypothesis, is impossible). Thus, the assumption that $W_0 = 0$ leads to a contradiction, which completes the proof of the theorem.

Theorem II is very important; it shows that knowing any two solutions of equation (**) which satisfy the conditions of the theorem we can construct the general solution and, hence, find particular solutions satisfying any given initial conditions.

II. Nonhomogeneous Linear Equations. Consider a nonhomogeneous linear equation of the second order

$$y'' + a_1 y' + a_2 y = f(x) \quad (*)$$

We shall say that the homogeneous equation

$$y'' + a_1 y' + a_2 y = 0 \quad (**)$$

obtained from the given equation (*) by replacing the nonhomogeneity term $f(x)$ by zero, *corresponds* to equation (*). We shall prove the following theorem describing the structure of the general solution of the nonhomogeneous equation (*).

Theorem. The *general solution* of the nonhomogeneous equation (*) can be written as a sum of the general solution of the corresponding homogeneous equation (**) and a particular solution of the given equation (*).

Proof. Let $\Phi(x)$ designate the general solution of equation (**) and $\varphi(x)$ an arbitrary particular solution of equation (*). Take the function

$$y = \Phi(x) + \varphi(x)$$

and differentiate it twice:

$$y' = \Phi'(x) + \varphi'(x), \quad y'' = \Phi''(x) + \varphi''(x)$$

Substituting the expressions of y , y' , and y'' into the left-hand side of the given equation (*) we obtain

$$\begin{aligned} & \Phi''(x) + \varphi''(x) + a_1 [\Phi'(x) + \varphi'(x)] + a_2 [\Phi(x) + \varphi(x)] = \\ & = [\Phi''(x) + a_1 \Phi'(x) + a_2 \Phi(x)] + [\varphi''(x) + a_1 \varphi'(x) + a_2 \varphi(x)] \end{aligned}$$

The expression in the first square brackets is equal to zero because $\Phi(x)$ is the solution of the homogeneous equation (**), and the expression in the second square brackets is equal to $f(x)$ since $\varphi(x)$ satisfies the nonhomogeneous equation (*). Consequently, the function $y = \Phi(x) + \varphi(x)$ is in fact a solution of equation (*). According to Theorem II, Sec. I, this function can be written in the form

$$y = C_1 y_1 + C_2 y_2 + \varphi(x)$$

where y_1 and y_2 are particular solutions of the corresponding homogeneous equation and $\varphi(x)$ a particular solution of the nonhomogeneous equation.

To complete the proof let the reader write down the system of equations for determining C_1 and C_2 for given initial conditions and show that it always has a unique solution.

170. Homogeneous Linear Differential Equations of the Second Order with Constant Coefficients. The problem of solving linear differential equations of the second order with variable coefficients a_1 and a_2 is extremely complicated and we shall not treat it here. Let us study linear equations with *constant* coefficients

$$y'' + a_1 y' + a_2 y = f(x) \quad (*)$$

where a_1 and a_2 are constant numbers. As before, we begin with homogeneous equations.

Consider a homogeneous linear differential equation of the second order

$$y'' + a_1 y' + a_2 y = 0 \quad (**)$$

where a_1 and a_2 are constant. Our aim is to construct the *general solution* of this equation.

Let us try to construct a function of the form $y = e^{rx}$ satisfying equation (**) where r is a real or complex number. On analysing the simple calculations given below the reader will see that there is every reason to expect that for a certain r the function e^{rx} is a solution of equation (**).

We have

$$y' = re^{rx}, \quad y'' = r^2 e^{rx}$$

and, consequently, there must be the identity

$$e^{rx} (r^2 + a_1 r + a_2) = 0$$

Since $e^{rx} \neq 0$, it follows that

$$r^2 + a_1 r + a_2 = 0 \quad (A)$$

We thus see that the function e^{rx} is a solution of equation (**) if r is a root of quadratic equation (A).

Equation (A) is called the *characteristic equation* and r a *characteristic root*.

It is seen that in order to form the characteristic equation we must replace in the given equation (**) y by unity and each derivative of the unknown function (that is y' and y'') by r to the power equal to the order of the derivative (i.e., by r and r^2 , respectively).

Here, in connection with the roots r_1 and r_2 of the characteristic equation, we should distinguish between the following three

cases (the coefficients a_1 and a_2 are always supposed to be real numbers):

- (1) r_1 and r_2 are real and distinct: $r_1 \neq r_2$;
- (2) r_1 and r_2 are real and their values coincide: $r_1 = r_2$ (this means that r_1 is a two-fold root of equation (A));
- (3) r_1 and r_2 are conjugate complex numbers: $r_1 = \alpha + \beta i$ and $r_2 = \alpha - \beta i$, $\beta \neq 0$.

Let us separately analyse each case.

(1) *The roots of the characteristic equation are real and distinct: $r_1 \neq r_2$.* In this case either root can be taken as the exponent r in the function e^{rx} , and thus we readily obtain two solutions of equation (**): $e^{r_1 x}$ and $e^{r_2 x}$. It is clear that their ratio is not a constant quantity: $\frac{e^{r_2 x}}{e^{r_1 x}} = e^{(r_2 - r_1)x} \neq \text{const.}$

If the roots r_1 and r_2 of the characteristic equation are real and distinct the general solution is given by the formula

$$y = C_1 e^{r_1 x} + C_2 e^{r_2 x}$$

where C_1 and C_2 are arbitrary constants.

The determinant W_0 of the system of linear algebraic equations from which, for given initial conditions at a point x_0 , the values of C_1 and C_2 are found (see Sec. 169, I), is expressed by the formula

$$W_0 = \begin{vmatrix} e^{r_1 x_0} & e^{r_2 x_0} \\ r_1 e^{r_1 x_0} & r_2 e^{r_2 x_0} \end{vmatrix} = e^{(r_1 + r_2)x_0} (r_2 - r_1)$$

Since $r_2 \neq r_1$, we have $W_0 \neq 0$ for all the values of x_0 , which is in accordance with the theorem on the general solution of the homogeneous linear differential equation.

Example. Let us solve the equation

$$y'' - y' - 2y = 0$$

Write the characteristic equation:

$$r^2 - r - 2 = 0$$

Its roots are

$$r_1 = 2 \quad \text{and} \quad r_2 = -1$$

Hence, the general solution is

$$y = C_1 e^{2x} + C_2 e^{-x}$$

Now let us set the initial conditions $y|_{x=0} = 2$, $y'|_{x=0} = -5$ and determine the corresponding particular solution. Here the algebraic system for C_1 and C_2 has the form

$$\begin{aligned} C_1 + C_2 &= 2 \\ 2C_1 - C_2 &= -5 \end{aligned}$$

whence $C_1 = -1$, $C_2 = 3$, and the sought-for particular solution is found:

$$y = -e^{2x} + 3e^{-x}$$

(2) *The roots of the characteristic equation are real and coincide:* $r_1 = r_2$. In this case the above procedure only yields one solution $y_1 = e^{r_1 x}$. We shall show that the function

$$y_2 = xe^{r_1 x}$$

can be taken as a second solution of the equation. Differentiate the function y_2 twice:

$$\begin{aligned} y_2' &= e^{r_1 x} + r_1 x e^{r_1 x}, \\ y_2'' &= 2r_1 e^{r_1 x} + r_1^2 x e^{r_1 x} \end{aligned}$$

Substitute these expressions into the left-hand side of equation (**):

$$\begin{aligned} 2r_1 e^{r_1 x} + r_1^2 x e^{r_1 x} + a_1 (e^{r_1 x} + r_1 x e^{r_1 x}) + a_2 x e^{r_1 x} &= \\ = e^{r_1 x} [x(r_1^2 + a_1 r_1 + a_2) + (2r_1 + a_1)] \end{aligned}$$

Since r_1 is a root of the characteristic equation we have $r_1^2 + a_1 r_1 + a_2 = 0$ and, since it is a two-fold root, *Vieta's theorem* tells that $r_1 + r_1 = -a_1$, that is $2r_1 + a_1 = 0$. Thus, the expression in square brackets is equal to zero, and the function $y_2 = xe^{r_1 x}$ is in fact a solution of equation (**).

Hence, in the case of a real two-fold root of the characteristic equation the general solution of equation (**) is

$$y = (C_1 + C_2 x) e^{r_1 x}$$

Here it is also readily checked that the determinant W_0 does not vanish for any value of x_0 :

$$\begin{vmatrix} e^{r_1 x_0} & x_0 e^{r_1 x_0} \\ r_1 e^{r_1 x_0} & e^{r_1 x_0} + r_1 x_0 e^{r_1 x_0} \end{vmatrix} = e^{2r_1 x_0} \neq 0$$

Example. Let us solve the equation

$$y'' - 6y' + 9y = 0$$

The characteristic equation

$$r^2 - 6r + 9 = 0$$

has one two-fold root

$$r_1 = r_2 = 3$$

and, consequently, the general solution is written in the form

$$y = (C_1 + C_2 x) e^{3x}$$

(3) *The roots of the characteristic equation are conjugate complex numbers:* $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$, $\beta \neq 0$. If we admit as solu-

tions complex functions of a real argument (see § 6 of Chapter IV¹), then, as before, the functions $e^{(\alpha+i\beta)x}$ and $e^{(\alpha-i\beta)x}$ are two solutions of equation (**), their ratio (which is equal to $e^{2i\beta x}$) being nonconstant. The general solution can be written as

$$C_1 e^{(\alpha+i\beta)x} + C_2 e^{(\alpha-i\beta)x}$$

where C_1 and C_2 are arbitrary complex constants. Let the reader apply the differentiation rules for complex functions to check that this expression is a solution of the equation.

To obtain the real solution we shall use the following simple property: if equation (**) with real coefficients has a complex solution $y = u(x) + iv(x)$ each of the functions $u(x)$ and $v(x)$ is a solution of the equation. Indeed, on differentiating the function y and substituting the result into the equation we obtain

$$(u'' + iv'') + a_1(u' + iv') + a_2(u + iv) = 0$$

Regrouping the summands we write

$$(u'' + a_1 u' + a_2 u) + i(v'' + a_1 v' + a_2 v) = 0$$

Since a complex expression is equal to zero if and only if its real and imaginary parts are zero, there must be

$$u'' + a_1 u' + a_2 u = 0 \quad \text{and} \quad v'' + a_1 v' + a_2 v = 0$$

which means that the functions $u(x)$ and $v(x)$ are solutions.

Since, by Euler's formula (see Sec. 74), we have

$$e^{(\alpha+i\beta)x} = e^{\alpha x} e^{i\beta x} = e^{\alpha x} (\cos \beta x + i \sin \beta x) = e^{\alpha x} \cos \beta x + i e^{\alpha x} \sin \beta x$$

the above property shows that the functions $e^{\alpha x} \cos \beta x$ and $e^{\alpha x} \sin \beta x$ are solutions to equation (**); their ratio is clearly distinct from a constant. Now, knowing the two particular solutions we readily construct the general solution

$$y = C_1 e^{\alpha x} \cos \beta x + C_2 e^{\alpha x} \sin \beta x = e^{\alpha x} (C_1 \cos \beta x + C_2 \sin \beta x)$$

where C_1 and C_2 are no longer complex but *real* arbitrary constants. In this construction we have not used the other complex solution $e^{(\alpha-i\beta)x}$; if we take its real part $e^{\alpha x} \cos \beta x$ and imaginary part $-e^{\alpha x} \sin \beta x$ to construct the general solution the final result is obviously the same as before.

Thus, if the characteristic equation has two conjugate complex roots the general solution is

$$y = e^{\alpha x} (C_1 \cos \beta x + C_2 \sin \beta x)$$

¹ If this section has not yet been studied the reader may omit the computations and limit himself to memorizing the ultimate conclusion given in bold face type and to considering the given examples.

Let the reader verify that in this case as well the determinant W_0 (see Sec. 169, I) is different from zero.

Example. $y'' - 4y' + 13y = 0$

Write down the characteristic equation:

$$r^2 - 4r + 13 = 0$$

Its roots are

$$r_1 = 2 + 3i \quad \text{and} \quad r_2 = 2 - 3i$$

Therefore the general solution is given by the formula

$$y = e^{2x} (C_1 \cos 3x + C_2 \sin 3x)$$

The theory of vibrations often deals with equations in which $a_1 = 0$ and $a_2 > 0$. Let us write such an equation in the form

$$y'' + \omega^2 y = 0$$

where ω is a positive constant.

From the characteristic equation $r^2 + \omega^2 = 0$ we find

$$r_1 = \omega i, \quad r_2 = -\omega i$$

Therefore

$$y = C_1 \cos \omega x + C_2 \sin \omega x$$

This general solution can be written in a different form if we introduce two new arbitrary constants A and φ connected with C_1 and C_2 by the relations

$$C_1 = A \sin \varphi, \quad C_2 = A \cos \varphi$$

which are equivalent to $A = \sqrt{C_1^2 + C_2^2}$ and $\tan \varphi = \frac{C_1}{C_2}$. Then

$$y = A (\sin \varphi \cos \omega x + \cos \varphi \sin \omega x) = A \sin (\omega x + \varphi)$$

Setting arbitrary initial conditions we can find uniquely determined values of the constants C_1 and C_2 or A and φ .

Let us consider the case of *boundary* conditions (see the end of Sec. 165) when the situation may be different. For definiteness, let $\omega = 1$, then the general solution is $y = C_1 \cos x + C_2 \sin x$. Take the boundary conditions $y|_{x=0} = 0$, $y|_{x=\pi} = y_0$. The first condition implies that $C_1 = 0$; then the second condition becomes $C_2 \sin \pi = y_0$ and shows that if $y_0 \neq 0$ it is impossible to determine C_2 , which means that there is no solution satisfying the given boundary conditions. But if $y_0 = 0$, any function $y = C_2 \sin x$ satisfies both the equation and the boundary conditions. Let the reader show that if the second condition is taken at any point $x \neq \pi$ the problem possesses a unique solution.

171. Nonhomogeneous Linear Differential Equations of the Second Order with Constant Coefficients. The Method of Complex Amplitudes.

1. Now we consider a linear equation of the form

$$y'' + a_1 y' + a_2 y = f(x) \quad (*)$$

with constant coefficients a_1 and a_2 and nonzero nonhomogeneity term $f(x)$.

As is known (see Sec. 169, II), the general solution of this equation is the sum of the general solution of the corresponding homogeneous equation and a particular solution of equation (*).

Since it is already known how the general solution of the homogeneous equation is found, it only remains to determine a particular solution of the given equation (*). We shall begin with some special cases in which the solution can be found with the aid of the method of undetermined coefficients and then proceed to the general method.

1. Let the right-hand side of equation (*) be of the form

$$f(x) = P(x) e^{px}$$

where $P(x)$ is a polynomial. Then equation (*) possesses a particular solution

$$y = x^k Q(x) e^{px}$$

where $Q(x)$ is a polynomial of the same degree as $P(x)$; if the number p is not a root of the characteristic equation $r^2 + a_1 r + a_2 = 0$ then $k = 0$ and if it is a root then k is equal to the multiplicity of that root.

Assuming that there is a solution of this form we then find the coefficients of the polynomial $Q(x)$ with the aid of the method of undetermined coefficients.

This rule remains valid when $p = 0$, that is when the right-hand side only involves a polynomial; in this case it is necessary to check whether the number 0 is a characteristic root and to determine its multiplicity provided it is a root. In particular cases $P(x)$ can be a polynomial of zeroth degree, that is a constant.

Below are some examples intended to help the reader to understand this method of determining a particular solution.

Examples. (1) $y'' - 2y' + y = 1 + x$, $y|_{x=0} = 2$, $y'|_{x=0} = -3$.

Here the characteristic equation

$$r^2 - 2r + 1 = 0$$

has the two-fold root $r = 1$. Hence, the general solution of the homogeneous equation is

$$y = (C_1 + C_2 x) e^x$$

The right-hand side of the given equation has the form indicated above with $p=0$ and $P(x)=1+x$. Since the number 0 is not a root of the characteristic equation the particular solution is sought in the form

$$y = Ax + B$$

where A and B are constants to be found.

Differentiating this expression and substituting into the differential equation we obtain

$$-2A + Ax + B = 1 + x$$

Equating the coefficients in like powers of x on both sides of the equality we derive the relations

$$A = 1, \quad -2A + B = 1$$

whence $A=1$ and $B=3$. Thus, the function

$$y = x + 3$$

is a *particular* solution of the given equation, and its *general solution is the function*

$$y = (C_1 + C_2 x) e^x + (x + 3)$$

On finding the general solution of the equation we proceed to determine its particular solution corresponding to the given initial conditions¹. To this end we find

$$y' = (C_1 + C_2 x) e^x + C_2 e^x + 1$$

Then

$$2 = C_1 + 3, \quad -3 = C_1 + C_2 + 1$$

whence $C_1 = -1$ and $C_2 = -3$. Thus, the desired solution is the function

$$y = -(1 + 3x) e^x + x + 3$$

$$(2) \quad y'' - 4y' + 3y = 3e^{2x}.$$

Here the characteristic equation is

$$r^2 - 4r + 3 = 0$$

its roots being $r_1 = 1$ and $r_2 = 3$. Consequently, the general solution of the corresponding homogeneous equation is

$$y = C_1 e^x + C_2 e^{3x}$$

The right-hand side of the given equation is of the form discussed above with $P(x)=3$. Thus, here we have a polynomial of

¹ We have already found one particular solution $y=x+3$ but it does not satisfy the initial conditions. This solution has been used for constructing the general solution of the equation from which we now find another particular solution corresponding to the initial conditions.

zeroth degree; the number $p=2$ is not a root of the characteristic equation and therefore $k=0$. The particular solution is therefore taken in the form

$$y = Ae^{2x}$$

where A is a constant that should be determined.

We have $y' = 2Ae^{2x}$ and $y'' = 4Ae^{2x}$. The substitution of y , y' and y'' into the equation results in

$$4Ae^{2x} - 8Ae^{2x} + 3Ae^{2x} = 3e^{2x}$$

whence $A = -3$. Consequently, the sought-for particular solution is the function $y = -3e^{2x}$ and the general solution is

$$y = C_1e^x + C_2e^{3x} - 3e^{2x}$$

$$(3) \quad y'' - 4y' + 3y = xe^x.$$

Here the left-hand side is the same as in Example 2. Therefore the general solution of the corresponding homogeneous equation is already known:

$$y = C_1e^x + C_2e^{3x}$$

Since $p=1$, that is p is a simple (one-fold) root of the characteristic equation, we have $k=1$. The polynomial $P(x)=x$ is of the first degree, and therefore the form of the desired particular solution is

$$y = x(Ax + B)e^x = (Ax^2 + Bx)e^x$$

We now differentiate it twice:

$$\begin{aligned} y' &= (2Ax + B)e^x + (Ax^2 + Bx)e^x \\ y'' &= 2Ae^x + 2(2Ax + B)e^x + (Ax^2 + Bx)e^x \end{aligned}$$

On substituting these expressions into the equation we obtain

$$2Ae^x + 2(2Ax + B)e^x + (Ax^2 + Bx)e^x - 4[(2Ax + B)e^x + (Ax^2 + Bx)e^x] + 3(Ax^2 + Bx)e^x = e^x(-4Ax + 2A - 2B) = xe^x$$

whence

$$2A - 2B = 0, \quad -4A = 1$$

that is

$$A = -\frac{1}{4}, \quad B = -\frac{1}{4}$$

Thus, we have determined the particular solution $y = -\frac{1}{4}(x^2 + x)e^x$ and hence found the general one:

$$y = C_1e^x + C_2e^{3x} - \frac{1}{4}(x^2 + x)e^x$$

2. Let the right-hand side of equation (*) be of the form

$$f(x) = a \cos qx + b \sin qx$$

If the numbers $\pm iq$ are not characteristic roots, the equation possesses the particular solution

$$x = A \cos qx + B \sin qx$$

If the numbers $\pm iq$ satisfy the characteristic equation there is a particular solution of the form

$$y = x (A \cos qx + B \sin qx)$$

In particular cases when $a=0$ or $b=0$ the solution should nevertheless be sought in the indicated general form involving both $\cos qx$ and $\sin qx$.

Examples. (1) $y'' + 4y' + 13y = 5 \sin 2x$.

The characteristic equation $r^2 + 4r + 13 = 0$ has the roots $r = -2 \pm 3i$. Hence, the general solution of the corresponding homogeneous equation is

$$y = e^{-2x} (C_1 \cos 3x + C_2 \sin 3x)$$

Since the numbers $\pm 2i$ are not roots of the characteristic equation, we look for a particular solution of the form

$$y = A \cos 2x + B \sin 2x$$

Differentiating twice we obtain

$$y' = -2A \sin 2x + 2B \cos 2x$$

$$y'' = -4A \cos 2x - 4B \sin 2x$$

The substitution into the equation yields

$$\begin{aligned} -4A \cos 2x - 4B \sin 2x - 8A \sin 2x + 8B \cos 2x + 13A \cos 2x + \\ + 13B \sin 2x = 5 \sin 2x \end{aligned}$$

Equating the coefficients in $\sin 2x$ and $\cos 2x$ on both sides of the equality (the coefficient in $\cos 2x$ on the right-hand side is equal to zero) we obtain

$$-8A + 9B = 5, \quad 9A + 8B = 0$$

whence $A = -\frac{8}{29}$ and $B = \frac{9}{29}$. Hence, the particular solution is the function $-\frac{8}{29} \cos 2x + \frac{9}{29} \sin 2x$, and the general solution is

$$y = e^{-2x} (C_1 \cos 3x + C_2 \sin 3x) - \frac{8}{29} \cos 2x + \frac{9}{29} \sin 2x$$

(2) $y'' + \omega^2 y = a \sin nx$.

The general solution of the corresponding homogeneous equation has already been found on page 611:

$$y = C_1 \cos \omega x + C_2 \sin \omega x$$

If $n \neq \omega$ a particular solution of the given equation is sought in the form

$$y = A \cos nx + B \sin nx$$

Differentiating this expression and substituting the result into the equation (let the reader carry out the calculations) we find

$$A(\omega^2 - n^2) \cos nx + B(\omega^2 - n^2) \sin nx = a \sin nx$$

whence

$$A = 0 \quad \text{and} \quad B = \frac{a}{\omega^2 - n^2}$$

Therefore the general solution of the nonhomogeneous equation is

$$y = C_1 \cos \omega x + C_2 \sin \omega x + \frac{a}{\omega^2 - n^2} \sin nx$$

If $n = \omega$ the above form of the particular solution does not apply. In this case it should be sought in the form

$$y = x(A \cos \omega x + B \sin \omega x)$$

We have

$$y' = (A \cos \omega x + B \sin \omega x) + x\omega(-A \sin \omega x + B \cos \omega x)$$

and

$$y'' = 2\omega(-A \sin \omega x + B \cos \omega x) - x\omega^2(A \cos \omega x + B \sin \omega x)$$

On substituting into the equation we receive

$$2\omega(-A \sin \omega x + B \cos \omega x) - x\omega^2(A \cos \omega x + B \sin \omega x) + x\omega^2(A \cos \omega x + B \sin \omega x) = a \sin \omega x$$

that is

$$2\omega(-A \sin \omega x + B \cos \omega x) = a \sin \omega x$$

whence

$$A = -\frac{a}{2\omega}, \quad B = 0$$

Thus, the function

$$y = -\frac{a}{2\omega} x \cos \omega x$$

is a particular solution of the equation for $n = \omega$.

Consequently, the general solution is given by

$$y = C_1 \cos \omega x + C_2 \sin \omega x - \frac{a}{2\omega} x \cos \omega x$$

Now we pass to a more general form of nonhomogeneity terms to which the method of undetermined coefficients is also applicable.

3. If the right-hand side of equation (*) has the form

$$f(x) = e^{px} [P_1(x) \cos qx + P_2(x) \sin qx]$$

where $P_1(x)$ and $P_2(x)$ are polynomials and the numbers $p \pm iq$ are not roots of the characteristic equation, there exists a particular solution of the form

$$y = e^{px} [R_1(x) \cos qx + R_2(x) \sin qx]$$

where $R_1(x)$ and $R_2(x)$ are polynomials of degree equal to the highest degree of the polynomials $P_1(x)$ and $P_2(x)$.

If the numbers $p \pm iq$ are characteristic roots this expression for the particular solution should be additionally multiplied by x .

Case 1 is obtained from this general expression for $q=0$, and Case 2 for $p=0$, $P_1(x)=a$ and $P_2(x)=b$ (the constants are regarded as polynomials of zeroth degree).

It should be noted that one of the polynomials $R_1(x)$ and $R_2(x)$ may happen to be of degree less than the one taken originally, that is some of the leading coefficients of that polynomial may be equal to zero.

Example. $y'' + y = 4x \sin x$.

Here $p=0$, $q=1$, and the numbers $\pm i$ are the roots of the characteristic equation $r^2 + 1 = 0$. Therefore the particular solution is taken in the form

$$y = x [(Ax + B) \cos x + (A_1x + B_1) \sin x]$$

We have

$$y'' = [-Ax^2 + (4A_1 - B)x + (2A + 2B_1)] \cos x + \\ + [-A_1x^2 - (4A + B_1)x + (2A_1 - 2B)] \sin x$$

The substitution into the equation results in

$$[2A_1x + (A + B_1)] \cos x + [-2Ax + (A_1 - B)] \sin x = 2x \sin x$$

This equality becomes an identity only if

$$2A_1 = 0, \quad A + B_1 = 0, \quad -2A = 2 \quad \text{and} \quad A_1 - B = 0$$

whence

$$A = -1, \quad B = 0, \quad A_1 = 0, \quad B_1 = 1$$

Hence, the sought-for particular solution is

$$y = x (\sin x - x \cos x)$$

The general solution is given by the formula

$$y = C_1 \cos x + C_2 \sin x + x (\sin x - x \cos x)$$

The following remark is of practical importance:

Let the right-hand side of the equation $y'' + a_1y' + a_2y = f(x)$ be equal to the sum of two functions: $f(x) = f_1(x) + f_2(x)$; if y_1 and y_2 are solutions of the equations with the same left-hand side and the right-hand sides equal, respectively, to $f_1(x)$ and $f_2(x)$, the function $y_1 + y_2$ is a solution of the given equation.

For we have

$$(y_1'' + y_2'') + a_1(y_1' + y_2') + a_2(y_1 + y_2) = (y_1'' + a_1y_1' + a_2y_1) + (y_2'' + a_1y_2' + a_2y_2) = f_1(x) + f_2(x) = f(x)$$

This means that if we manage to find solutions of the equations with right-hand sides equal to the constituents the given nonhomogeneity term is formed of, the sought-for solution is readily found as the sum of these solutions.

For instance, according to this remark, the equation

$$y'' - 4y' + 3y = 3e^{2x} + xe^x$$

(see Examples (2) and (3) on pages 613, 614) has a particular solution of the form

$$y = -3e^{2x} - \frac{1}{4}(x^2 + x)e^x$$

Consequently, its general solution is

$$y = C_1e^x + C_2e^{3x} - 3e^{2x} - \frac{1}{4}(x^2 + x)e^x$$

II. The Method of Complex Amplitudes. In all the examples considered by now the right-hand side of the equation is a real function. However, in the theory of vibrations we are often interested in equations whose nonhomogeneity terms are *complex* functions of a real argument. The role of the argument is usually played by time which we shall denote as t . Denoting the unknown function by $z(t)$, we have, in the new notation, the equation

$$\ddot{z} + a_1\dot{z} + a_2z = f(t) \quad (*)$$

It is natural to expect that if $f(t) = f_1(t) + if_2(t)$ is a complex function the solution is also a complex function: $z(t) = x(t) + iy(t)$.

In what follows the coefficients a_1 and a_2 are assumed to be *real* numbers since this is usually the case in applications. The substitution of the expressions of $z(t)$ and of its derivatives into equation (*) yields

$$(\ddot{x} + i\ddot{y}) + a_1(\dot{x} + i\dot{y}) + a_2(x + iy) = f_1(t) + if_2(t)$$

On separating the real and imaginary parts we arrive at the two equalities

$$\ddot{x} + a_1\dot{x} + a_2x = f_1(t), \quad \ddot{y} + a_1\dot{y} + a_2y = f_2(t) \quad (**)$$

We thus see that after a particular solution of equation (*) has been found we can separate its real and imaginary parts and thus obtain particular solutions of equations (**). The converse is also true: any particular solutions of equations (**) obviously generate a particular solution of equation (*).

Let us dwell in more detail on the case when the right member of equation (*) is an exponential function: $f(t) = Ce^{i\omega t}$. The complex quantity $C = re^{i\alpha}$ where $r = |C|$ and $\alpha = \arg C$ is called the *complex amplitude*. Since

$$Ce^{i\omega t} = re^{i(\omega t + \alpha)} = r \cos(\omega t + \alpha) + ir \sin(\omega t + \alpha)$$

the determination of a particular solution to the equation

$$\ddot{z} + a_1 \dot{z} + a_2 z = Ce^{i\omega t} \quad (***)$$

enables us to find particular solutions to equations (**) with right-hand members $r \cos(\omega t + \alpha)$ and $r \sin(\omega t + \alpha)$. The latter differential equations are often dealt with in electrotechnical problems. The replacement of trigonometric functions by exponential ones simplifies the calculations; this technique is referred to as the *method of complex amplitudes*.

Let us look for a particular solution of equation (***) of the form $z = Se^{i\omega t}$ where S is a complex number. On differentiating, substituting into the equation and cancelling by $e^{i\omega t}$ we get

$$(-\omega^2 + a_1 i\omega + a_2)S = C \quad \text{whence} \quad S = \frac{C}{a_2 - \omega^2 + ia_1\omega}$$

(it is supposed that the denominator of S is nonzero). Let us denote the modulus of the complex number S by ρ and the argument by β . Then $S = \rho e^{i\beta}$, and

$$z = Se^{i\omega t} = \rho e^{i(\omega t + \beta)} = \rho \cos(\omega t + \beta) + i\rho \sin(\omega t + \beta)$$

Hence, the equation $\ddot{x} + a_1 \dot{x} + a_2 x = r \cos(\omega t + \beta)$ possesses the particular solution $\rho \cos(\omega t + \beta)$; similarly, the equation with right member $r \sin(\omega t + \alpha)$ has the solution $\rho \sin(\omega t + \beta)$.

172. The Method of Variation of Arbitrary Constants¹. Let us proceed to the general method for finding a particular solution of a nonhomogeneous equation

$$y'' + a_1 y' + a_2 y = f(x) \quad (*)$$

where $f(x)$ is an *arbitrary* continuous function.

To apply this method we must know the general solution of the corresponding homogeneous equation. The method of variation of arbitrary constants is applicable both to equations with constant coefficients and to equations whose coefficients a_1 and a_2 are functions of x . But, since practically, in the general case, we are only able to solve homogeneous equations with constant coefficients, the method we are going to discuss here will be used for solving equations of that kind.

¹ This method was inaugurated by J.L. Lagrange.

Let the homogeneous equation

$$y'' + a_1 y' + a_2 y = 0 \quad (**)$$

corresponding to equation (*) have the general solution

$$y = C_1 y_1 + C_2 y_2$$

where C_1 and C_2 are arbitrary constants.

We shall try to find a solution of equation (*) having the form

$$y = C_1(x) y_1 + C_2(x) y_2 \quad (***)$$

where $C_1(x)$ and $C_2(x)$ are some unknown functions which should be determined, and y_1 and y_2 two known particular solutions of homogeneous equation (**). For brevity, we shall write C_1 and C_2 instead of $C_1(x)$ and $C_2(x)$ but it should be borne in mind that these quantities are dependent on x . Here we have to determine two functions C_1 and C_2 which must be chosen so that equation (*) is satisfied. This gives us one relation for C_1 and C_2 . But, generally, to determine two unknown quantities we must have two relations, and therefore a second relation can be introduced arbitrarily. Differentiate equality (***):

$$y' = C_1' y_1 + C_2' y_2 + C_1 y_1' + C_2 y_2'$$

It turns out that it is most convenient to impose as a second condition on C_1 and C_2 the requirement that the expression for y' should have the same form as in the case of constant C_1 and C_2 .

To this end we put

$$C_1' y_1 + C_2' y_2 = 0 \quad (A)$$

Then

$$y' = C_1 y_1' + C_2 y_2'$$

Now we find the second derivative of y :

$$y'' = C_1' y_1' + C_1 y_1'' + C_2' y_2' + C_2 y_2''$$

On substituting y , y' and y'' into the left-hand side of equation (*) we obtain

$$\begin{aligned} C_1' y_1 + C_1 y_1'' + C_2' y_2 + C_2 y_2'' + a_1 C_1 y_1' + a_1 C_2 y_2' + a_2 C_1 y_1 + a_2 C_2 y_2 = \\ = C_1' y_1 + C_2' y_2 + C_1 (y_1'' + a_1 y_1' + a_2 y_1) + C_2 (y_2'' + a_1 y_2' + a_2 y_2) = f(x) \end{aligned}$$

The expressions in both parentheses are equal to zero since y_1 and y_2 are solutions to the homogeneous equation (**). Hence, for the function $y = C_1 y_1 + C_2 y_2$ to satisfy equation (*) the condition

$$C_1' y_1 + C_2' y_2 = f(x) \quad (B)$$

should hold. Conditions (A) and (B) result in the system of

equations

$$\begin{aligned}C_1' y_1 + C_2' y_2 &= 0 \\ C_1' y_1' + C_2' y_2' &= f(x)\end{aligned}$$

As was indicated in Sec. 169, I, the determinant of this system does not vanish at any point x :

$$\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} \neq 0$$

Therefore we can first find C_1' and C_2' and then, integrating, determine the functions C_1 and C_2 themselves. If, when integrating C_1' and C_2' , we introduce the corresponding arbitrary constants, we directly obtain the general solution of the given nonhomogeneous equation.

Example. Let us solve the equation

$$y'' + y = \tan x$$

The characteristic equation of the corresponding homogeneous equation $y'' + y = 0$ is $r^2 + 1 = 0$. Its roots are $r = \pm i$. Therefore $y_1 = \cos x$ and $y_2 = \sin x$.

Taking the solution of the given equation in the form

$$y = C_1 \cos x + C_2 \sin x$$

we obtain the system of equations

$$\begin{aligned}C_1' \cos x + C_2' \sin x &= 0 \\ -C_1' \sin x + C_2' \cos x &= \tan x\end{aligned}$$

for finding C_1 and C_2 . On solving the system we receive

$$C_1' = -\frac{\sin^2 x}{\cos x}, \quad C_2' = \sin x$$

The integration results in

$$C_1 = -\int \frac{\sin^2 x}{\cos x} dx = \int \left(\cos x - \frac{1}{\cos x} \right) dx = \sin x - \ln \tan \left(\frac{\pi}{4} + \frac{x}{2} \right) + k_1$$

(see Formula 69 in the table of integrals) and

$$C_2 = \int \sin x dx = -\cos x + k_2$$

where k_1 and k_2 are arbitrary constants. Finally, we write down the general solution of the given equation:

$$\begin{aligned}y &= \left[\sin x - \ln \tan \left(\frac{\pi}{4} + \frac{x}{2} \right) + k_1 \right] \cos x + [-\cos x + k_2] \sin x = \\ &= k_1 \cos x + k_2 \sin x - \cos x \ln \tan \left(\frac{\pi}{4} + \frac{x}{2} \right)\end{aligned}$$

173. **Linear Differential Equations of the n th Order.** The theory of linear differential equations of the second order presented in the foregoing sections is readily extended to the linear differential equations of the n th order ($n > 2$).

The general form of a linear equation of the n th order is

$$y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + \dots + a_{n-1} y' + a_n y = f(x) \quad (*)$$

where the coefficients $a_1, a_2, \dots, a_{n-1}, a_n$ and the right member $f(x)$ are continuous functions of the independent variable x in a given finite or infinite interval. In particular, the coefficients can be constant.

As before, for the given *nonhomogeneous* equation we introduce the corresponding *homogeneous* equation (without right member):

$$y^{(n)} + a_1 y^{(n-1)} + a_2 y^{(n-2)} + \dots + a_{n-1} y' + a_n y = 0 \quad (**)$$

To state the fundamental theorem on the structure of the general solution of such an equation we need the notion of a *linearly independent* system of functions.

Consider a system of functions $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ defined in a common interval. Remember that a *linear combination* of the functions is an expression of the form

$$C_1 \varphi_1(x) + C_2 \varphi_2(x) + \dots + C_n \varphi_n(x)$$

where $C_1, C_2, \dots, C_n$ are constant coefficients.

Definition. A *system of functions* $\varphi_1(x), \varphi_2(x), \dots, \varphi_n(x)$ is said to be *linearly independent* if none of the functions is representable as a linear combination of the other functions.

For instance, this definition implies that there cannot be an equality

$$\varphi_1(x) = k_2 \varphi_2(x) + \dots + k_n \varphi_n(x)$$

with constant coefficients $k_2, \dots, k_n$.

It follows that none of the linearly independent functions can be identically equal to zero. For instance, if we supposed that $\varphi_1(x) \equiv 0$, the above equality would be fulfilled for $k_2 = \dots = k_n = 0$.

In particular, two functions $\varphi_1(x)$ and $\varphi_2(x)$ are linearly independent if their ratio is not a constant: $\frac{\varphi_2(x)}{\varphi_1(x)} \neq \text{const.}$

A system of functions which is not linearly independent is said to be *linearly dependent*. For example, the system

$$\varphi_1(x) = x, \quad \varphi_2(x) = x^2, \quad \varphi_3(x) = x^3, \quad \varphi_4(x) = 2x - x^2$$

is linearly dependent. Indeed, the function $\varphi_4(x)$ is a linear combination of the other functions:

$$\varphi_4(x) = 2\varphi_1(x) - \varphi_2(x)$$

Note that it is not required that a linear combination expressing $\varphi_n(x)$ in terms of $\varphi_j(x)$, $j=1, 2, \dots, n-1$, should involve *all* the functions $\varphi_1(x), \dots, \varphi_{n-1}(x)$; some of the coefficients in the functions $\varphi_1(x), \dots, \varphi_{n-1}(x)$ (and even all of them!) can be equal to zero.

Now we proceed to state the theorem on the structure of the general solution of equation (**).

Theorem. If $y_1, y_2, \dots, y_n$ are n linearly independent particular solutions of equation (**) the general solution of the equation is their linear combination

$$y = C_1 y_1 + C_2 y_2 + \dots + C_n y_n \quad (***)$$

with n arbitrary constant coefficients $C_1, C_2, \dots, C_n$.

If the solutions $y_1, y_2, \dots, y_n$ are *linearly dependent*, at least one of them is expressible linearly in terms of the other $n-1$ solutions, and then expression (***) involves in fact less than n independent arbitrary constants and therefore does not provide the general solution.

There is a simple condition guaranteeing the linear independence of particular solutions $y_1, y_2, \dots, y_n$. Namely, this condition requires that a determinant called the *Wronskian*¹ of the functions $y_1, y_2, \dots, y_n$ should be nonzero:

$$W(y_1, y_2, \dots, y_n) = \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \vdots & \vdots & \dots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix} \neq 0$$

This determinant involves the functions $y_1, \dots, y_n$ and their derivatives up to order $n-1$ inclusive.

The inequality sign in this relation is understood in the sense that the Wronskian does not vanish for any value of x . The general proof of this condition is not presented here; for the case $n=2$ it is given in Sec. 169, I. The Wronskian plays an important role in finding the particular solution satisfying given initial conditions.

For, if we are given initial conditions

$$y|_{x=x_0} = y_0, \quad y'|_{x=x_0} = y_0', \quad \dots, \quad y^{(n-1)}|_{x=x_0} = y_0^{(n-1)}$$

then in order to isolate the desired particular solution from the general solution

$$y = C_1 y_1 + C_2 y_2 + \dots + C_n y_n$$

¹ Named after J.M. Hönené-Wronski (1778-1853), a Polish mathematician.

particular solutions of the form

$$\begin{aligned} e^{\alpha x} \cos \beta x, \quad x e^{\alpha x} \cos \beta x, \quad \dots, \quad x^{t-1} e^{\alpha x} \cos \beta x \\ e^{\alpha x} \sin \beta x, \quad x e^{\alpha x} \sin \beta x, \quad \dots, \quad x^{t-1} e^{\alpha x} \sin \beta x \end{aligned}$$

The total sum of the multiplicities of all the roots is equal to the degree n of the characteristic equation; therefore the total number of the above particular solutions coincides with the order of the differential equation.

Here we do not present the proof of the fact that these particular solutions are linearly independent, which means that they constitute a fundamental system of solutions.

The general solution of the given differential equation is a linear combination of these particular solutions with arbitrary constant coefficients.

Example. $y^{(5)} + y^{(4)} + 2y''' + 2y'' + y' + y = 0$

As is readily seen, the characteristic equation

$$r^5 + r^4 + 2r^3 + 2r^2 + r + 1 = 0$$

has $r = -1$ as root; on dividing the equation by $r + 1$ we obtain

$$r^4 + 2r^2 + 1 = 0$$

that is

$$(r^2 + 1)^2 = 0$$

Hence, we have

$$r_1 = -1, \quad r_2 = r_3 = i, \quad r_4 = r_5 = -i$$

Consequently, the general solution of the differential equation is

$$y = C_1 e^{-x} + (C_2 + C_3 x) \cos x + (C_4 + C_5 x) \sin x$$

A particular solution of the nonhomogeneous equation

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} y' + a_n y = f(x)$$

where $f(x)$ has the form

$$f(x) = e^{px} [P_1(x) \cos qx + P_2(x) \sin qx]$$

($P_1(x)$ and $P_2(x)$ are polynomials in x) is found according to the same rules as in the case of nonhomogeneous equations of the second order (see Sec. 171).

It turns out that the equation possesses a particular solution of the form

$$y = x^k e^{px} [R_1(x) \cos qx + R_2(x) \sin qx]$$

where $R_1(x)$ and $R_2(x)$ are polynomials of the highest degree of the polynomials $P_1(x)$ and $P_2(x)$, and k is the multiplicity of $p \pm qi$ as roots of the characteristic equation. If $p \pm qi$ are not roots of the characteristic equation the number k is put to be equal

to zero. If the right member of the equation $f(x)$ is not a function of the indicated type one should apply the method of variation of arbitrary constants.

In conclusion we note that in many cases it is preferable to solve linear differential equations with constant coefficients with the aid of the methods of the so-called *operational calculus*. These methods are based on the theory of functions of a complex variable and are widely used in the theory of automatic control and in the electrotechnical problems. On these questions see, for instance, [3] and [19].

175. Vibrations. Resonance.

I. Mechanical Vibrations. Vibration problems play an important role in modern engineering and physics. There are many cases when the phenomenon of *vibrations* (synonymously, *oscillations*) is described by linear differential equations of the second order (linear vibrations) which in the simplest cases have constant coefficients. We shall begin with *mechanical vibrations*.

Suppose that a moving body (for simplicity, of unit mass) is under the action of a force directed toward the state of equilibrium, the magnitude of the force being proportional to the deviation from that state. If the distance from the body to the state of equilibrium is denoted by s and the proportionality coefficient (the "stiffness factor") by ω^2 the magnitude of the force is equal to $\omega^2 s$ (as will be seen, ω is the *natural frequency* of free vibrations in the absence of the resistance of the medium). Such a force is said to be *restoring*.

Furthermore, let the motion be in a medium in which the body is subjected to a force of resistance. We shall suppose that this force is proportional to the velocity of motion (its direction is opposite to the motion). According to this hypothesis, the magnitude of this force is equal to $2ks$ if $2k$ designates the proportionality factor. Thus, this is a *resisting force*, $k > 0$ being called the *resistance coefficient*.

Finally, let the body be also acted upon by an external force whose magnitude is given as a function $f(t)$ of time t . This is a *disturbing (exciting) force*.

Now let us write down the differential equation of motion. To find the resultant of all the forces we should obviously take the restoring force and the resisting force with the minus sign since the direction of the former is opposite to the direction in which s is reckoned, and the latter is opposite to the velocity. We thus come to the equation

$$\ddot{s} = -2k\dot{s} - \omega^2 s + f(t)$$

that is

$$\ddot{s} + 2k\dot{s} + \omega^2 s = f(t) \quad (*)$$

Hence, the phenomenon of mechanical vibrations is described by a linear differential equation of the second order with constant coefficients.

If there is no exciting force, i.e. $f(t) \equiv 0$, the vibrations are spoken of as *free* or *natural*; the free oscillations are described by the homogeneous equation corresponding to equation (*). If the exciting force is nonzero then, in addition to free vibrations, there appear *forced* vibrations, and the motion is described by the general nonhomogeneous equation (*).

A restoring force can be *elastic*. For instance, the case of an elastic restoring force is realized as follows. Let a load be suspended by a vertical spring, its weight being balanced by the elastic force of the spring. If the load is given a vertical displacement from the equilibrium state it then goes into oscillations which are free if there is no additional force acting upon the load or upon the whole system or forced if there is such a force (for instance, this force can make the point of suspension of the load move according to a certain law). The counteracting force of the spring (the restoring force) is directly proportional to the deflection from the equilibrium position (i.e. to the deformation) only if the deformation is sufficiently small (*Hooke's law*). The proportionality factor is called the *spring constant* or the *spring stiffness* (or, as was mentioned, the *stiffness factor*).

II. Investigating Free Vibrations. Let us begin with investigating free vibrations described by the homogeneous linear differential equation of the second order with constant coefficients

$$\ddot{s} + 2k\dot{s} + \omega^2 s = 0$$

Its characteristic equation $r^2 + 2kr + \omega^2 = 0$ has the roots $-k \pm \sqrt{k^2 - \omega^2}$. The following three cases are possible here.

(1) *The resistance coefficient is greater than the natural frequency ω : $k > \omega$.*

Then the solution is

$$s = C_1 e^{-\delta_1 t} + C_2 e^{-\delta_2 t}$$

where

$$\delta_1 = k - \sqrt{k^2 - \omega^2} > 0 \quad \text{and} \quad \delta_2 = k + \sqrt{k^2 - \omega^2} > 0$$

We see that the process does not involve any vibration; as t increases the deflection s decreases and tends to zero for $t \rightarrow \infty$, the system approaching the equilibrium state (which is practically achieved during a finite time). As we say, in this case there is an *aperiodic (damped) motion*. From the physical point of view this is accounted for by the fact that the action of the decelerative resisting force is considerably stronger than that of the resto-

ring force (responsible for vibrations), and therefore the motion is damped out before the system returns to the equilibrium position.

(2) *The resistance coefficient is equal to the frequency ω : $k = \omega$.*
Then the solution is

$$s = e^{-kt} (C_1 + C_2 t).$$

This case does not differ, in principle, from the preceding one in the sense that beginning with a certain value of t the quantity s becomes a decreasing function tending to zero for $t \rightarrow \infty$.

(3) *The resistance coefficient is less than the frequency ω : $k < \omega$.*
Then the solution is written in the form

$$s = e^{-kt} (C_1 \cos k_1 t + C_2 \sin k_1 t) \quad \text{where } k_1 = \sqrt{\omega^2 - k^2}$$

Making the transformation indicated on page 611 we obtain

$$s = Ae^{-kt} \sin(k_1 t + \varphi_0)$$

where A and φ_0 are arbitrary constants. To specify their values one must set initial conditions, that is the initial deviation of the body from the equilibrium position and its initial velocity.

In this case the body is in a vibration motion which is called *damped harmonic oscillation*.

If the resistance coefficient is equal to zero ($k=0$) we have

$$s = A \sin(\omega t + \varphi_0)$$

that is an ordinary (undamped) harmonic oscillation (as was mentioned, its *natural frequency* is equal to ω).

If $k \neq 0$ the amplitude of the vibration is Ae^{-kt} , and, in contrast to the case of pure harmonic vibrations, it is not a constant quantity but depends on time and tends to zero as $t \rightarrow \infty$. Consequently, the body gradually approaches the equilibrium position vibrating about it, its amplitude tending to zero according to an exponential law. By analogy with the case of pure harmonic vibrations, the quantity $k_1 = \sqrt{\omega^2 - k^2}$ is called the *natural frequency* of the damped harmonic vibration, $T = \frac{2\pi}{k_1}$ is called the *period* of the vibration and φ_0 the *initial phase*. The magnitude Ae^{-kt} is spoken of as the *amplitude* of the damped vibrations. The logarithm of the amplitude $\ln A - kt$ decreases with the constant rate k . To compute the maximum deviation s we find the derivative

$$\dot{s} = -kAe^{-kt} \sin(k_1 t + \varphi_0) + Ak_1 e^{-kt} \cos(k_1 t + \varphi_0)$$

(which is the velocity of motion) and equate it to zero. This results in the equation

$$\tan(k_1 t + \varphi_0) = \frac{k_1}{k}$$

for determining the extremal values of t . It is clear that the maxima and the minima of the deviation (which have opposite signs) alternate, the time interval between them being equal to the half-period $\frac{T}{2} = \frac{\pi}{k_1}$. The ratio of two consecutive maxima¹ is equal to

$$\Delta = \frac{Ae^{-kt_0}}{Ae^{-k(t_0+T)}} = e^{kT}.$$

The constant quantity $\ln \Delta = kT$ is called the *logarithmic decrement*.

The graph of the damped harmonic oscillation $s = Ae^{-kt} \sin k_1 t$ (for $\varphi_0 = 0$) is schematically depicted in Fig. 244. The construction

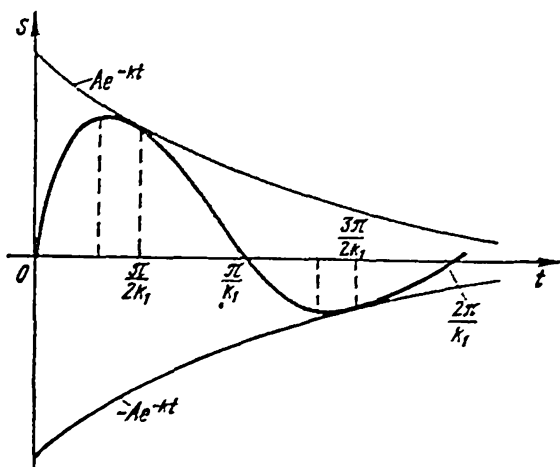


Fig. 244

of the graph takes into account the fact that at the common points of the curves Ae^{-kt} and $Ae^{-kt} \sin k_1 t$, that is for $k_1 t = \frac{\pi}{2} + 2n\pi$, the derivatives of these functions are equal (let the reader show this), and therefore the curves have common tangents at these points. The same applies to the curves $-Ae^{-kt}$ and $Ae^{-kt} \sin k_1 t$. As has been already mentioned, the time interval between two consecutive points of maximum and minimum is equal to the half-period, i.e. to $\frac{\pi}{k_1}$.

III. Investigating Forced Vibrations. Resonance. Now we proceed to study forced oscillations described by the nonhomo-

¹ Here t_0 is one of the values of t for which s attains a maximum. Practically, the quantity Δ is usually determined by measuring two consecutive maximum amplitudes and finding their ratio although it is apparent that the ratio of any two values of s separated by a time interval T is also equal to Δ .

neous equation

$$\ddot{s} + 2k\dot{s} + \omega^2 s = f(t)$$

We shall consider the case when the resistance coefficient is less than the natural frequency of undamped oscillations ($k < \omega$) and the exciting force is periodic ($f(t) = a \sin \omega_1 t$), which is most important for practical problems.

We shall first find a particular solution of the nonhomogeneous equation

$$\ddot{s} + 2k\dot{s} + \omega^2 s = a \sin \omega_1 t$$

Let $k \neq 0$; then $\omega_1 i$ is not a root of the characteristic equation, and the solution should be sought for in the form $M \cos \omega_1 t + N \sin \omega_1 t$. On determining the coefficients M and N as was indicated in the foregoing sections we arrive at the solution

$$s = A_1 \sin(\omega_1 t + \varphi_1)$$

where

$$A_1 = \sqrt{M^2 + N^2} = \frac{a}{\sqrt{4k^2\omega_1^2 + (\omega^2 - \omega_1^2)^2}}$$

and

$$\tan \varphi_1 = \frac{M}{N} = \frac{-2k\omega_1}{\omega^2 - \omega_1^2}$$

Let the reader carry out all the calculations or derive this solution with the aid of the method of complex amplitudes (see Sec. 171, II).

The particular solution found above describes a *forced vibration*. This is an undamped harmonic vibration with amplitude A_1 , frequency ω_1 , and initial phase φ_1 ; the quantity φ_1 characterizes the phase shift of the forced vibration relative to the phase of disturbing force.

The general form of vibrations described by the equation in question is

$$s = Ae^{-kt} \sin(k_1 t + \varphi_0) + A_1 \sin(\omega_1 t + \varphi_1)$$

This is the sum of natural (free) vibration and the forced vibration. The rate of damping of free vibration is usually high, and therefore, after a certain time, it can be practically neglected. Then the vibration reduces to the forced one:

$$s \approx A_1 \sin(\omega_1 t + \varphi_1)$$

Let us consider an interesting phenomenon characteristic of forced vibrations which is called *resonance*.

Suppose that the quantities k and ω are constant, that is the natural frequency k_1 of damped oscillations of the system is con-

stant. Let us take the expression for A_1 and rewrite it as

$$A_1 = \frac{a}{\omega^2 \sqrt{\frac{4k^2}{\omega^2} \left(\frac{\omega_1}{\omega}\right)^2 + \left[1 - \left(\frac{\omega_1}{\omega}\right)^2\right]^2}} = \frac{a}{\omega^2 \sqrt{\alpha^2 q^2 + (1 - q^2)^2}}$$

where $\alpha = \frac{2k}{\omega}$ and $q = \frac{\omega_1}{\omega}$. The amplitude A_1 is proportional to the quantity

$$\lambda = \lambda(q) = \frac{1}{\sqrt{\alpha^2 q^2 + (1 - q^2)^2}}$$

The character of the variation of λ as function of q is seen in Fig. 245 where the graph of the function $\lambda(q)$ is shown for

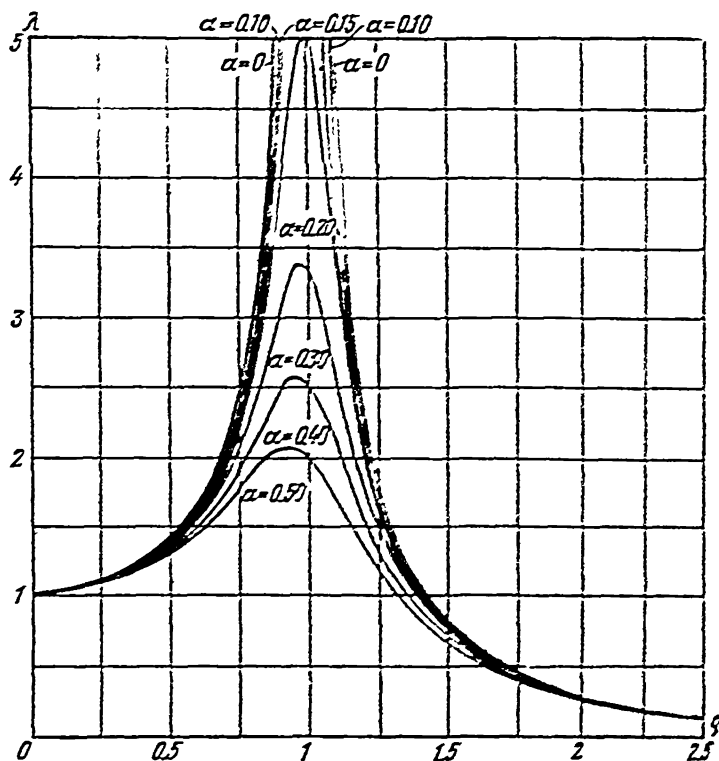


Fig. 245

$\alpha = 0.50, \alpha = 0.40, \alpha = 0.30, \alpha = 0.20, \alpha = 0.15, \alpha = 0.10$ and $\alpha = 0$. As a rule, the quantity α is rather small since the value of the resistance coefficient k is usually small relative to ω .

The investigation shows that when q increases, that is the excitation frequency ω_1 grows, the quantity λ and, together with it,

the amplitude of the *steady-state* forced vibration first increase until a maximum is attained and then rapidly decrease and tend to zero for $q \rightarrow \infty$ (that is $\omega_1 \rightarrow \infty$). The maximum is attained for $q = \sqrt{1 - \frac{\alpha^2}{2}}$, i.e. for $\omega_1 = \bar{\omega}_1 = \sqrt{\omega^2 - 2k^2}$. It is readily found that the maximum amplitude is $A_{1\max} = \frac{a}{2k\sqrt{\omega^2 - k^2}}$.

When the frequency of the disturbing force is close to the frequency $\bar{\omega}_1$ and k is small (this means that $\bar{\omega}_1 \approx \omega$) then, after some time passes, the amplitude of the harmonic vibrations of the body becomes very large compared with the amplitude of the sinusoidal force producing the vibration of the system. This phenomenon of an extensive growth of the amplitude of a vibration under the action of a small external excitation is known as *resonance*. The curves shown in Fig. 245 are called *resonance curves*.

Resonance plays an important role in engineering and physics. Every elastic body (for example, any construction) possesses its own natural frequency of vibrations solely dependent on the properties of the body. Suppose that under the action of an external force the body leaves the equilibrium position. If the frequency of the exciting force is close to the natural frequency, the action of the force, no matter how small it is, may produce a very strong and even destructive effect. When designing various constructions (machines, bridges, ships, airplanes, etc.) engineers should take into account the aspects of strength and durability connected with resonance. Resonance accounts for the phenomenon well known from experience when a small periodic excitation of an elastic body (of a bridge, for instance) leads to breaking.

If there is no resistance, that is $k=0$ (which is hardly possible in practice) the amplitude of the forced vibration is equal to $\frac{a}{|\omega^2 - \omega_1^2|}$ and turns into infinity when the excitation frequency coincides with the natural frequency of undamped harmonic oscillations (i.e. $\omega_1 = \omega$; see the curve corresponding to $\alpha=0$ in Fig. 245). If $\omega_1 \neq \omega$ the solution of the differential equation $\ddot{s} + \omega^2 s = a \sin \omega_1 t$ describing the vibrations is obtained from the general solution given above by putting $k=0$. But in the case $\omega_1 = \omega$ this does not apply since ωi then becomes a root of the characteristic equation $r^2 + \omega^2 = 0$. As was shown in Example (2) on pp. 615, 616, in this case the solution to the equation is the function

$$s = A \sin(\omega t + \varphi_0) - \frac{a}{2\omega} t \cos \omega t$$

The second summand (the so-called *secular term*) shows that the amplitude of the vibration increases indefinitely as $t \rightarrow \infty$, which again characterizes resonance.

IV. Vibrations in Electric Circuits. Electric processes in a circuit consisting of a resistance R , an inductance L and a capacitance C are completely analogous to mechanical vibrations. A similar question was discussed in Sec. 161 but there we did not take into account the capacitance. In the general case it is necessary to include the voltage drop across the capacitance. As is known

from physics, it is equal to $\frac{1}{C} \int_0^t i dt$ where i is the electric current flow.

We thus arrive at the equation

$$Li' + Ri + \frac{1}{C} \int_0^t i dt = V(t)$$

where $V(t)$ is an external voltage applied to the terminals of the circuit. The differentiation with respect to t results in the linear differential equation of the second order with constant coefficients

$$i'' + \frac{R}{L} i' + \frac{1}{LC} i = \frac{1}{L} V'(t)$$

This equation describes forced electric vibrations and is completely analogous to the equation of mechanical oscillations.

If the voltage V is constant (or is not applied at all, i.e. $V \equiv 0$) the equation is homogeneous. It specifies an electric current which must be damped since the resistance is nonzero ($R \neq 0$). Extending the analogy between mechanical and electrical oscillations we can say that (on condition that $\frac{R}{2L} < \sqrt{\frac{1}{CL}}$) this current corresponds to natural (free) vibrations in the circuit¹.

If a sinusoidal voltage is applied to the terminals there arise forced vibrations in the circuit. After some time (usually very short) the natural vibrations are damped out and do not affect the process which becomes a *steady-state* forced sinusoidal vibration.

The process lasting from the moment the voltage is switched on until a steady-state vibration arises is referred to as a *transient process*².

¹ In the case $\frac{R}{2L} \geq \sqrt{\frac{1}{CL}}$ there is no vibration in the circuit, and the current flow is aperiodic and damped.

² Here we speak of a particular case of a transient process. Its general definition can be found in electrical engineering text-books.

We shall not go into particulars related to resonance phenomenon in electric circuits since its mathematical aspects are similar to the case of mechanical vibrations.

In electrical engineering the method of analysing electric circuits described here is spoken of as *classical*. The investigation of transient processes is more often performed with the help of the operational calculus (the *operational method*) mentioned at the end of Sec. 174.

§ 4. Systems of Differential Equations

176. General Definitions. Normal Form of a System of Differential Equations. We encounter systems of differential equations when investigating processes described by more than one function. We have already dealt with some examples of this kind. For instance, in Sec. 151 we mentioned that the determination of vector lines of a field is connected with solving a system of differential equations. In footnote on p. 598 it was indicated that the solution of dynamical problems connected with studying curvilinear motion leads to a system of three differential equations in which the unknown functions are the projections of the radius-vector of the moving point on the coordinate axes and the independent variable is time. Later on we shall see that the electrotechnical problem of investigating two electric circuits with electromagnetic coupling is connected with the solution of a system of two differential equations. There are many other examples of this kind.

We shall begin with some basic definitions. Everywhere in this section we shall designate the independent variable by the letter t and the unknown functions dependent on that variable by $x_1(t)$, $x_2(t)$, ..., $x_n(t)$ or, if their number does not exceed three, by $x(t)$, $y(t)$ and $z(t)$.

Definition. A *system of differential equations* is a collection of equations each of which may involve the independent variable, the unknown functions and their derivatives.

It is always assumed that the number of the equations is equal to the number of the unknown functions. Here are examples of systems of differential equations (the derivatives with respect to t are marked by dots instead of primes):

$$\left. \begin{aligned} \dot{x} &= 2x + y + t + 1 \\ \dot{y} &= 3x - 2y + 5t \end{aligned} \right\} \begin{cases} t\ddot{x}_1 + \dot{x}_2 + 2x_1x_3 = 0 \\ 2\ddot{x}_2 + \ddot{x}_3 - 2tx_1 = 0 \\ \dot{x}_3 + 2x_1 + tx_2 = 0 \end{cases}$$

Any collection of functions

$$x_1 = x_1(t), \quad x_2 = x_2(t), \quad \dots, \quad x_n = x_n(t)$$

and introduce the two new auxiliary functions

$$y = \dot{x} \text{ and } z = \dot{y} = \ddot{x}$$

The given equation is then replaced by the equivalent system of three equations

$$\dot{x} = y, \quad \dot{y} = z, \quad \dot{z} = f(t, x, y, z)$$

which is a particular case of the general normal system (*). It appears obvious that the same technique can be applied to an equation of an arbitrary order n ; in this case the number of the auxiliary functions is $n-1$.

In the cases most often dealt with the converse is also true: Generally speaking, *a system of differential equations in normal form can be replaced by an equivalent differential equation whose order is equal to the number of the equations in the system.*

As an example, consider the system of equations

$$\dot{x} = y, \quad \dot{y} = z, \quad \dot{z} = x - y + z$$

Differentiate the first equation with respect to the variable t and replace the derivative $\dot{y}$ by its expression given by the second equation: $\ddot{x} = \dot{y} = z$. Now differentiate the latter equation repeatedly and replace the derivative $\dot{z}$ by its expression taken from the third equation: $\ddot{\ddot{x}} = \dot{z} = x - y + z$. Since $y = \dot{x}$ and $z = \ddot{x}$ we finally obtain $\ddot{\ddot{x}} = x - \dot{x} + \ddot{x}$, that is

$$\ddot{\ddot{x}} - \ddot{x} + \dot{x} - x = 0$$

which is a linear differential equation of the third order with constant coefficients.

Note that when eliminating the functions y and z we express them in terms of the function x and its derivatives. On finding the general solution of the resultant differential equation of the third order we obtain the expression for the function x involving three arbitrary constants. The remaining unknown functions are then found without integration from their expressions in terms of the function x already determined. Hence, the total number of the arbitrary constants is not increased and coincides with the order of the system, that is with the number of the equations.

Let us solve the equation of the third order obtained above. The corresponding characteristic equation $r^3 - r^2 + r - 1 \equiv (r^2 + 1)(r - 1) = 0$ has the roots $r_1 = 1$ and $r_{2,3} = \pm i$. Consequently,

$$x = C_1 e^t + C_2 \cos t + C_3 \sin t$$

For systems of differential equations in normal form we can state the theorem on the existence and uniqueness of a particular solution analogous to the theorems formulated in the foregoing sections for equations of first, second and higher orders:

Theorem. If the right-hand sides of a normal system are continuous together with their partial derivatives in a neighbourhood of a point $(t_0, x_{10}, x_{20}, \dots, x_{n0})$ then in a sufficiently small interval $[t_0 - h, t_0 + h]$ there exists a uniquely determined system of functions $x_1(t), x_2(t), \dots, x_n(t)$ representing a particular solution of the system which satisfies the given initial conditions $x_1|_{t=t_0} = x_{10}, x_2|_{t=t_0} = x_{20}, \dots, x_n|_{t=t_0} = x_{n0}$.

177*. Geometrical and Mechanical Interpretation of Solutions of a System of Differential Equations. Phase Space. For the sake of simplicity, let us begin with the normal system of two differential equations

$$\dot{x} = f_1(t, x, y), \quad \dot{y} = f_2(t, x, y) \quad (*)$$

Independently of the physical meaning of the variables x and y let us interpret them as the coordinates of a point in the xy -plane. The plane itself will be called the *phase plane* of the system. Every solution to the system

$$x = x(t), \quad y = y(t) \quad (**)$$

can be regarded as giving a parametric representation of a curve in the phase plane. Using mechanical terminology and assuming that t is time we can say that the functions $x(t)$ and $y(t)$ describe some laws of motion of the projections of the moving point on the coordinate axes while the curve (**) itself is the *trajectory* of motion. The derivatives $\dot{x}$ and $\dot{y}$ then represent the projections of the velocity of the moving point on the coordinate axes.

It should be stressed that all that has been said is only a convenient geometrical or mechanical illustration. In concrete problems the variables x and y may designate some quantities which are in no connection with the motion of a point; they can, for instance, describe electric current in electric circuits. Also note that in this interpretation a solution $x = x(t)$ of a single differential equation $\dot{x} = f(t, x)$ expresses a law of motion of a point along the axis of abscissas; then to various particular solutions there correspond different laws of motions. The interval of the x -axis in which the point moves may be common for all the particular solutions (if, for instance, the general solution of the equation is $x = \sin Ct$) or different for different solutions (if, for example, $x = C \sin t$). In the first of these simple examples we have harmonic oscillations with the same amplitude but different

frequencies while in the second example the frequency is constant and the amplitude may vary.

Let us come back to the system of two differential equations (*) and set initial conditions: $x|_{t=t_0}=\alpha$, $y|_{t=t_0}=\beta$. They specify a particular solution $x_0(t)$, $y_0(t)$ determining a law of motion along a trajectory, L_0 . Now let us change the initial conditions by taking another initial time moment t_1 and retaining the initial values of the phase coordinates x and y : $x|_{t=t_1}=\alpha$, $y|_{t=t_1}=\beta$. We then obtain a new solution $x_1(t)$, $y_1(t)$ and a new trajectory L_1 which, of course, may differ from L_0 . For, on finding from equation (*) the values of the derivatives $\dot{x}$ and $\dot{y}$ at the initial time moments t_0 and t_1 we may obtain different results. This means that if the motion starts from the point (α, β) at time moment t_0 we have a certain initial velocity vector which may change if the motion begins at another time moment t_1 ; in these two motions the points may start moving in different directions. Thus, through a fixed point of the phase plane there may pass different trajectories (*phase trajectories*). Consider a simple example.

(1) $\dot{x}=1$, $\dot{y}=2t$. These two equations are not connected with each other, and therefore we immediately obtain $x=t+C_1$, $y=t^2+C_2$. For the initial conditions $x|_{t=0}=0$, $y|_{t=0}=0$ the solution is expressed by the functions $x=t$ and $y=t^2$, the corresponding phase trajectory being the parabola $y=x^2$. But if we take the same initial values of x and y for another initial time moment $t=1$ the new solution is $x=t-1$, $y=t^2-1$. On eliminating t we obtain the equation of the new phase trajectory: $y=(x+1)^2-1$. It is also a parabola but not coincident with the former.

In applications the most important case is when the right-hand sides of equations (*) do not involve explicitly the independent variable t . Such a system of differential equations is said to be *autonomous*; it is written in the form

$$\frac{dx}{dt}=f_1(x, y), \quad \frac{dy}{dt}=f_2(x, y) \quad (***)$$

If the motion is described by equations of this type the velocity vector at every fixed point (α, β) is independent of time. It can therefore be proved that all the motions starting at the fixed point (α, β) (which can be arbitrary) at different time moments have one and the same trajectory (we do not present the proof here but the mechanical meaning of this property is quite evident). Imagine that we are watching the motion of a small ball rolling along a trajectory; then it is obvious that similar balls issued from one and the same point with the same velocity but at

different moments of time move in the same trajectory without impacting and overtaking each other (but there can of course be closed trajectories, for instance, circles or ellipses).

It can also be shown that for an autonomous system through each point of the phase plane there is only a single trajectory provided that the conditions of the existence theorem stated in Sec. 176 are fulfilled in a neighbourhood of that point. What has been said is demonstrated by the example below.

(2) $\frac{dx}{dt} = y, \frac{dy}{dt} = -x$. Eliminating the function y we get $\ddot{x} = \dot{y} = -x$, that is $\ddot{x} + x = 0$. Write the general solution of the latter equation in the form $x = A \sin(t + \varphi)$; then $y = A \cos(t + \varphi)$. The trajectories of the system are the circles $x^2 + y^2 = A^2$ with centre at the origin.

The equation of the family of the phase trajectories can be derived in another way by eliminating at the very beginning the differential of the independent variable dt which is readily achieved by dividing the second equation of the system by the first one: $\frac{dy}{dx} = -\frac{x}{y}$. Solving the first-order differential equation thus obtained we arrive at its general solution describing the same family of circles.

Let us briefly discuss one more question. Again consider system (***) and suppose that both right members turn into zero at a point (x_0, y_0) :

$$f_1(x_0, y_0) = 0, f_2(x_0, y_0) = 0$$

Then the functions $x = x_0$ and $y = y_0$ (simply speaking, the constants) are solutions to the system. This means that the ball placed at the point (x_0, y_0) remains motionless (in the state of rest) all the time. A point of this kind is spoken of as a *rest point* of the system. For instance, the origin is a rest point of system (2).

For the equation $\frac{dy}{dx} = \frac{f_2(x, y)}{f_1(x, y)}$ obtained from the original system by eliminating dt every rest point of the system is a *singular point* (see Sec. 164). It follows that the problem of investigating the structure of the family of integral curves in the vicinity of a singular point is closely related to the investigation of the properties of the trajectories of the corresponding autonomous system near the rest point. These questions are studied in the *stability theory* (e.g. see [3]).

All that has been said can be extended to systems of three and more equations without any significant changes. In the case $n=3$ we deal with three unknown functions $x(t)$, $y(t)$ and $z(t)$ and therefore consider phase trajectories in the three-dimen-

frequencies while in the second example the frequency is constant and the amplitude may vary.

Let us come back to the system of two differential equations (*) and set initial conditions: $x|_{t=t_0}=\alpha$, $y|_{t=t_0}=\beta$. They specify a particular solution $x_0(t)$, $y_0(t)$ determining a law of motion along a trajectory, L_0 . Now let us change the initial conditions by taking another initial time moment t_1 and retaining the initial values of the phase coordinates x and y : $x|_{t=t_1}=\alpha$, $y|_{t=t_1}=\beta$. We then obtain a new solution $x_1(t)$, $y_1(t)$ and a new trajectory L_1 which, of course, may differ from L_0 . For, on finding from equation (*) the values of the derivatives $\dot{x}$ and $\dot{y}$ at the initial time moments t_0 and t_1 we may obtain different results. This means that if the motion starts from the point (α, β) at time moment t_0 we have a certain initial velocity vector which may change if the motion begins at another time moment t_1 ; in these two motions the points may start moving in different directions. Thus, through a fixed point of the phase plane there may pass different trajectories (*phase trajectories*). Consider a simple example.

(1) $\dot{x}=1$, $\dot{y}=2t$. These two equations are not connected with each other, and therefore we immediately obtain $x=t+C_1$, $y=t^2+C_2$. For the initial conditions $x|_{t=0}=0$, $y|_{t=0}=0$ the solution is expressed by the functions $x=t$ and $y=t^2$, the corresponding phase trajectory being the parabola $y=x^2$. But if we take the same initial values of x and y for another initial time moment $t=1$ the new solution is $x=t-1$, $y=t^2-1$. On eliminating t we obtain the equation of the new phase trajectory: $y=(x+1)^2-1$. It is also a parabola but not coincident with the former.

In applications the most important case is when the right-hand sides of equations (*) do not involve explicitly the independent variable t . Such a system of differential equations is said to be *autonomous*; it is written in the form

$$\frac{dx}{dt} = f_1(x, y), \quad \frac{dy}{dt} = f_2(x, y) \quad (***)$$

If the motion is described by equations of this type the velocity vector at every fixed point (α, β) is independent of time. It can therefore be proved that all the motions starting at the fixed point (α, β) (which can be arbitrary) at different time moments have one and the same trajectory (we do not present the proof here but the mechanical meaning of this property is quite evident). Imagine that we are watching the motion of a small ball rolling along a trajectory; then it is obvious that similar balls issued from one and the same point with the same velocity but at

(2) If we know two solutions x_1, y_1, z_1 and x_2, y_2, z_2 the functions $x_1 + x_2, y_1 + y_2, z_1 + z_2$ are also a solution.

It follows that if we know three particular solutions (x_1, y_1, z_1) , (x_2, y_2, z_2) and (x_3, y_3, z_3) then the functions

$$\begin{aligned}x &= C_1 x_1 + C_2 x_2 + C_3 x_3 \\y &= C_1 y_1 + C_2 y_2 + C_3 y_3 \\z &= C_1 z_1 + C_2 z_2 + C_3 z_3\end{aligned}\quad (**)$$

represent a solution for any constants C_1, C_2 and C_3 . Let us find out when it is possible to obtain from this three-parameter family of solutions a particular solution satisfying any given initial conditions. If the initial conditions are

$$x|_{t=t_0} = x_0, \quad y|_{t=t_0} = y_0, \quad z|_{t=t_0} = z_0$$

the system of equations for determining the constants C_1, C_2 and C_3 has the form

$$\begin{aligned}C_1 x_{10} + C_2 x_{20} + C_3 x_{30} &= x_0 \\C_1 y_{10} + C_2 y_{20} + C_3 y_{30} &= y_0 \\C_1 z_{10} + C_2 z_{20} + C_3 z_{30} &= z_0\end{aligned}$$

where x_{i0}, y_{i0} and z_{i0} ($i = 1, 2, 3$) are, respectively, the values of the corresponding functions for $t = t_0$. For this system to have a unique solution for any initial conditions it is necessary and sufficient that the determinant

$$W = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

be nonzero for all the values of $t = t_0$. As before, any collection of three particular solutions satisfying this condition will be called a *fundamental system of solutions*. Expressions (**) in this case describe the *general solution* of the system.

Proceeding to the nonhomogeneous system

$$\begin{aligned}\dot{x} &= a_1(t)x + b_1(t)y + c_1(t)z + g_1(t) \\ \dot{y} &= a_2(t)x + b_2(t)y + c_2(t)z + g_2(t) \\ \dot{z} &= a_3(t)x + b_3(t)y + c_3(t)z + g_3(t)\end{aligned}\quad (***)$$

we state that *the general solution of the nonhomogeneous system (***) is the sum of the general solution of the corresponding homogeneous system (*) and a particular solution of the nonhomogeneous system.*

The proof of this assertion is similar to the proof of the analogous theorem for a single linear equation (see Sec. 169, II), and we leave it to the reader.

sional phase space $Oxyz$. Their geometrical study is of course more complicated than in the case of the plane. For $n > 3$ the visual geometric representation is no longer possible. But we generalize the geometrical terminology and speak of an n -dimensional phase space.

Now let us discuss a *hydrodynamical* interpretation of solutions of a system of differential equations. Consider a stationary velocity field of a flow of liquid $V = \{v_x, v_y, v_z\}$, and let $v_x = f_1(x, y, z)$, $v_y = f_2(x, y, z)$ and $v_z = f_3(x, y, z)$. The differential equations of the family of stream lines (flow lines; see Sec. 151) are written in the form

$$\frac{dx}{f_1(x, y, z)} = \frac{dy}{f_2(x, y, z)} = \frac{dz}{f_3(x, y, z)}$$

Introducing a parameter t and equating each of the above ratios to dt we arrive at the autonomous system of differential equations

$$\begin{aligned}\frac{dx}{dt} &= f_1(x, y, z), & \frac{dy}{dt} &= f_2(x, y, z), \\ \frac{dz}{dt} &= f_3(x, y, z)\end{aligned}$$

The solutions $x(t)$, $y(t)$, $z(t)$ of this system give us equations of the stream lines. These stream lines are independent of time, and the particles moving along different stream lines do not mix since the stream lines do not intersect. For a nonstationary motion of the liquid the structure of the flow becomes more complicated.

178. Systems of Linear Differential Equations. In applications we most often deal with systems of linear differential equations. For the sake of simplicity, we shall confine ourselves to considering a system of three equations with three unknown functions $x(t)$, $y(t)$ and $z(t)$. We shall take the system in its normal form and begin with studying the properties of the homogeneous system

$$\begin{aligned}\dot{x} &= a_1(t)x + b_1(t)y + c_1(t)z \\ \dot{y} &= a_2(t)x + b_2(t)y + c_2(t)z \\ \dot{z} &= a_3(t)x + b_3(t)y + c_3(t)z\end{aligned}\tag{*}$$

The functions $a_i(t)$, $b_i(t)$, $c_i(t)$ ($i = 1, 2, 3$) are supposed to be continuous.

Let the reader prove the following simple properties of this system:

(1) If a particular solution $x_1(t)$, $y_1(t)$, $z_1(t)$ of system of equations (*) is known the functions $Cx_1(t)$, $Cy_1(t)$, $Cz_1(t)$ where C is an arbitrary constant also form a solution.

(2) If we know two solutions x_1, y_1, z_1 and x_2, y_2, z_2 , the functions $x_1 + x_2, y_1 + y_2, z_1 + z_2$ are also a solution.

It follows that if we know three particular solutions (x_1, y_1, z_1) , (x_2, y_2, z_2) and (x_3, y_3, z_3) then the functions

$$\begin{aligned}x &= C_1 x_1 + C_2 x_2 + C_3 x_3 \\y &= C_1 y_1 + C_2 y_2 + C_3 y_3 \\z &= C_1 z_1 + C_2 z_2 + C_3 z_3\end{aligned}\quad (**)$$

represent a solution for any constants C_1, C_2 and C_3 . Let us find out when it is possible to obtain from this three-parameter family of solutions a particular solution satisfying any given initial conditions. If the initial conditions are

$$x|_{t=t_0} = x_0, \quad y|_{t=t_0} = y_0, \quad z|_{t=t_0} = z_0$$

the system of equations for determining the constants C_1, C_2 and C_3 has the form

$$\begin{aligned}C_1 x_{10} + C_2 x_{20} + C_3 x_{30} &= x_0 \\C_1 y_{10} + C_2 y_{20} + C_3 y_{30} &= y_0 \\C_1 z_{10} + C_2 z_{20} + C_3 z_{30} &= z_0\end{aligned}$$

where x_{i0}, y_{i0} and z_{i0} ($i = 1, 2, 3$) are, respectively, the values of the corresponding functions for $t = t_0$. For this system to have a unique solution for any initial conditions it is necessary and sufficient that the determinant

$$W = \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

be nonzero for all the values of $t = t_0$. As before, any collection of three particular solutions satisfying this condition will be called a *fundamental system of solutions*. Expressions (**) in this case describe the *general solution* of the system.

Proceeding to the nonhomogeneous system

$$\begin{aligned}\dot{x} &= a_1(t)x + b_1(t)y + c_1(t)z + g_1(t) \\ \dot{y} &= a_2(t)x + b_2(t)y + c_2(t)z + g_2(t) \\ \dot{z} &= a_3(t)x + b_3(t)y + c_3(t)z + g_3(t)\end{aligned}\quad (***)$$

we state that the *general solution of the nonhomogeneous system (***) is the sum of the general solution of the corresponding homogeneous system (*) and a particular solution of the nonhomogeneous system*.

The proof of this assertion is similar to the proof of the analogous theorem for a single linear equation (see Sec. 169, II), and we leave it to the reader.

Generally speaking, the problem of solving a system of linear differential equations is equivalent to that of solving one linear equation of the corresponding order¹ (in our case a third-order equation). Indeed, as was indicated, to reduce system (*) or (***) to a single third-order equation we should differentiate one of the equations of the system, for instance, the first. Replacing each time the derivatives of the unknown functions in the right-hand side by their expressions we see that all the consecutively computed derivatives of the function x are expressed linearly in terms of the functions x , y and z . Therefore, conversely, the functions y and z can be represented as linear expressions in x , $\dot{x}$ and $\ddot{x}$. Hence, the resultant differential equation is linear. It is obvious that if the original system of equations is homogeneous the resultant equation is also homogeneous while a nonhomogeneous system may lead to a nonhomogeneous or a homogeneous equation. The latter possibility is illustrated by the system $\dot{x} = x + y + \frac{t^2}{2}$, $\dot{y} = x - t$ which reduces to the equation $\ddot{x} - \dot{x} - x = 0$.

We did not dwell on solving differential equations with *variable* coefficients, and therefore we shall not discuss the methods of solving systems of such equations and shall proceed to systems of linear equations with *constant* coefficients.

179. Systems of Linear Differential Equations with Constant Coefficients². Consider a *homogeneous* system of the form

$$\begin{aligned}\dot{x} &= a_1x + b_1y + c_1z \\ \dot{y} &= a_2x + b_2y + c_2z \\ \dot{z} &= a_3x + b_3y + c_3z\end{aligned}\tag{*}$$

where a_i , b_i and c_i ($i = 1, 2, 3$) are constant. As was mentioned at the end of the foregoing section, this system is reducible to a single homogeneous linear equation of the third order. It is clear that, as in the example on p. 637, this always leads to an equation with constant coefficients. Since the structure of the solutions of such an equation has been already studied, in practical work it is advisable to look directly for particular solutions of the form

$$x = k_1 e^{rt}, \quad y = k_2 e^{rt}, \quad z = k_3 e^{rt}\tag{**}$$

(where k_1 , k_2 , k_3 and r are undetermined constants to be found) instead of reducing the system to one equation.

¹ The example on p. 638 shows that sometimes there may appear several higher-order equations; in this example each of the equations contains only one unknown function.

² The reader is advised to recall the properties of homogeneous linear algebraic equations before studying this section.

On differentiating the functions x , y and z and substituting the results into system (*) we obtain

$$\begin{aligned} k_1 r e^{rt} &= a_1 k_1 e^{rt} + b_1 k_2 e^{rt} + c_1 k_3 e^{rt} \\ k_2 r e^{rt} &= a_2 k_1 e^{rt} + b_2 k_2 e^{rt} + c_2 k_3 e^{rt} \\ k_3 r e^{rt} &= a_3 k_1 e^{rt} + b_3 k_2 e^{rt} + c_3 k_3 e^{rt} \end{aligned}$$

Cancelling by e^{rt} and transposing all the terms to one side we arrive at the system of homogeneous linear algebraic equations in k_1 , k_2 and k_3 :

$$\begin{aligned} (a_1 - r) k_1 + b_1 k_2 + c_1 k_3 &= 0 \\ a_2 k_1 + (b_2 - r) k_2 + c_2 k_3 &= 0 \\ a_3 k_1 + b_3 k_2 + (c_3 - r) k_3 &= 0 \end{aligned} \quad (***)$$

For this system of homogeneous linear equations to possess a *nonzero (nontrivial)* solutions (of course, we are only interested in such solutions) it is necessary and sufficient that its determinant be equal to zero:

$$\begin{vmatrix} a_1 - r & b_1 & c_1 \\ a_2 & b_2 - r & c_2 \\ a_3 & b_3 & c_3 - r \end{vmatrix} = 0$$

On writing this determinant explicitly we obtain an algebraic equation of the third degree with respect to r which is called the *characteristic equation* of system (*).

Thus, a solution of form (**) exists if and only if the number r is a root of the characteristic equation.

Let us first suppose that all the roots of the characteristic equation are *real* and *simple* (i.e. not multiple). If r_1 is one of the roots its substitution into system of equations (**) results in

$$\begin{aligned} (a_1 - r_1) k_1 + b_1 k_2 + c_1 k_3 &= 0 \\ a_2 k_1 + (b_2 - r_1) k_2 + c_2 k_3 &= 0 \\ a_3 k_1 + b_3 k_2 + (c_3 - r_1) k_3 &= 0 \end{aligned}$$

By the hypothesis, the determinant of the latter system is equal to zero. Assume without proof that if r_1 is a simple root of the characteristic equation (we only confine ourselves to such roots now), at least one of the minors of the second order of the determinant is different from zero. Then one of the equations is a consequence of the other two equations, and the system reduces to two equations in three unknowns. The general solution of such a system depends on one arbitrary parameter. If, for instance, the system reduces to the first two equations, the numbers

$$k_1^{(1)} = \begin{vmatrix} b_1 & c_1 \\ b_2 - r_1 & c_2 \end{vmatrix}, \quad k_2^{(1)} = \begin{vmatrix} c_1 & a_1 - r_1 \\ c_2 & a_2 \end{vmatrix}, \quad k_3^{(1)} = \begin{vmatrix} a_1 - r_1 & b_1 \\ a_3 & b_2 - r_1 \end{vmatrix}$$

Generally speaking, the problem of solving a system of linear differential equations is equivalent to that of solving one linear equation of the corresponding order¹ (in our case a third-order equation). Indeed, as was indicated, to reduce system (*) or (***) to a single third-order equation we should differentiate one of the equations of the system, for instance, the first. Replacing each time the derivatives of the unknown functions in the right-hand side by their expressions we see that all the consecutively computed derivatives of the function x are expressed linearly in terms of the functions x , y and z . Therefore, conversely, the functions y and z can be represented as linear expressions in x , $\dot{x}$ and $\ddot{x}$. Hence, the resultant differential equation is linear. It is obvious that if the original system of equations is homogeneous the resultant equation is also homogeneous while a nonhomogeneous system may lead to a nonhomogeneous or a homogeneous equation. The latter possibility is illustrated by the system $\dot{x} = x + y + \frac{t^2}{2}$, $\dot{y} = x - t$ which reduces to the equation $\ddot{x} - \dot{x} - x = 0$.

We did not dwell on solving differential equations with *variable* coefficients, and therefore we shall not discuss the methods of solving systems of such equations and shall proceed to systems of linear equations with *constant* coefficients.

179. Systems of Linear Differential Equations with Constant Coefficients². Consider a *homogeneous* system of the form

$$\begin{aligned}\dot{x} &= a_1x + b_1y + c_1z \\ \dot{y} &= a_2x + b_2y + c_2z \\ \dot{z} &= a_3x + b_3y + c_3z\end{aligned}\tag{*}$$

where a_i , b_i and c_i ($i = 1, 2, 3$) are constant. As was mentioned at the end of the foregoing section, this system is reducible to a single homogeneous linear equation of the third order. It is clear that, as in the example on p. 637, this always leads to an equation with constant coefficients. Since the structure of the solutions of such an equation has been already studied, in practical work it is advisable to look directly for particular solutions of the form

$$x = k_1 e^{rt}, \quad y = k_2 e^{rt}, \quad z = k_3 e^{rt}\tag{**}$$

(where k_1 , k_2 , k_3 and r are undetermined constants to be found) instead of reducing the system to one equation.

¹ The example on p. 638 shows that sometimes there may appear several higher-order equations; in this example each of the equations contains only one unknown function.

² The reader is advised to recall the properties of homogeneous linear algebraic equations before studying this section.

On differentiating the functions x , y and z and substituting the results into system (*) we obtain

$$\begin{aligned}k_1 r e^{rt} &= a_1 k_1 e^{rt} + b_1 k_2 e^{rt} + c_1 k_3 e^{rt} \\k_2 r e^{rt} &= a_2 k_1 e^{rt} + b_2 k_2 e^{rt} + c_2 k_3 e^{rt} \\k_3 r e^{rt} &= a_3 k_1 e^{rt} + b_3 k_2 e^{rt} + c_3 k_3 e^{rt}\end{aligned}$$

Cancelling by e^{rt} and transposing all the terms to one side we arrive at the system of homogeneous linear algebraic equations in k_1 , k_2 and k_3 :

$$\begin{aligned}(a_1 - r) k_1 + b_1 k_2 + c_1 k_3 &= 0 \\a_2 k_1 + (b_2 - r) k_2 + c_2 k_3 &= 0 \\a_3 k_1 + b_3 k_2 + (c_3 - r) k_3 &= 0\end{aligned} \quad (***)$$

For this system of homogeneous linear equations to possess a *nonzero (nontrivial)* solutions (of course, we are only interested in such solutions) it is necessary and sufficient that its determinant be equal to zero:

$$\begin{vmatrix} a_1 - r & b_1 & c_1 \\ a_2 & b_2 - r & c_2 \\ a_3 & b_3 & c_3 - r \end{vmatrix} = 0$$

On writing this determinant explicitly we obtain an algebraic equation of the third degree with respect to r which is called the *characteristic equation* of system (*).

Thus, a solution of form (**) exists if and only if the number r is a root of the characteristic equation.

Let us first suppose that all the roots of the characteristic equation are *real* and *simple* (i.e. not multiple). If r_1 is one of the roots its substitution into system of equations (**) results in

$$\begin{aligned}(a_1 - r_1) k_1 + b_1 k_2 + c_1 k_3 &= 0 \\a_2 k_1 + (b_2 - r_1) k_2 + c_2 k_3 &= 0 \\a_3 k_1 + b_3 k_2 + (c_3 - r_1) k_3 &= 0\end{aligned}$$

By the hypothesis, the determinant of the latter system is equal to zero. Assume without proof that if r_1 is a simple root of the characteristic equation (we only confine ourselves to such roots now), at least one of the minors of the second order of the determinant is different from zero. Then one of the equations is a consequence of the other two equations, and the system reduces to two equations in three unknowns. The general solution of such a system depends on one arbitrary parameter. If, for instance, the system reduces to the first two equations, the numbers

$$k_1^{(1)} = \begin{vmatrix} b_1 & c_1 \\ b_2 - r_1 & c_2 \end{vmatrix}, \quad k_2^{(1)} = \begin{vmatrix} c_1 & a_1 - r_1 \\ c_2 & a_2 \end{vmatrix}, \quad k_3^{(1)} = \begin{vmatrix} a_1 - r_1 & b_1 \\ a_3 & b_3 - r_1 \end{vmatrix}$$

form one of the solutions. (These formulas are readily obtained if we transpose the terms with k_3 to the right-hand side and solve the resultant system of nonhomogeneous equations.) All the other solutions are found by multiplying the numbers $k_1^{(1)}$, $k_1^{(2)}$ and $k_3^{(1)}$ by one and the same arbitrary constant. Operating in the same manner on all the roots of the characteristic equation we determine three systems of functions each of which is a solution to system (*).

These systems of functions are

$$\begin{array}{lll} k_1^{(1)}e^{r_1 t}, & k_2^{(1)}e^{r_1 t}, & k_3^{(1)}e^{r_1 t} \\ k_1^{(2)}e^{r_2 t}, & k_2^{(2)}e^{r_2 t}, & k_3^{(2)}e^{r_2 t} \\ k_1^{(3)}e^{r_3 t}, & k_2^{(3)}e^{r_3 t}, & k_3^{(3)}e^{r_3 t} \end{array}$$

Here we do not present the proof of the fact that these functions constitute a fundamental system of solutions. The general solution of system (*) is written in the form

$$\begin{aligned} x &= C_1 k_1^{(1)}e^{r_1 t} + C_2 k_1^{(2)}e^{r_2 t} + C_3 k_1^{(3)}e^{r_3 t} \\ y &= C_1 k_2^{(1)}e^{r_1 t} + C_2 k_2^{(2)}e^{r_2 t} + C_3 k_2^{(3)}e^{r_3 t} \\ z &= C_1 k_3^{(1)}e^{r_1 t} + C_2 k_3^{(2)}e^{r_2 t} + C_3 k_3^{(3)}e^{r_3 t} \end{aligned}$$

Examples. (1) Let us solve the system

$$\begin{aligned} \dot{x} &= z \\ \dot{y} &= -4x - y - 4z \\ \dot{z} &= -y \end{aligned}$$

In this case the system of equations (***) has the form

$$\begin{aligned} -rk_1 &+ k_3 = 0 \\ -4k_1 - (1+r)k_2 - 4k_3 &= 0 \\ -k_2 - rk_3 &= 0 \end{aligned}$$

Write down the characteristic equation:

$$\begin{vmatrix} -r & 0 & 1 \\ -4 & -(1+r) & -4 \\ 0 & -1 & -r \end{vmatrix} = -r^3 - r^2 + 4r + 4 = (4-r^2)(r+1) = 0$$

Its roots are $r_1 = -1$, $r_2 = -2$ and $r_3 = 2$. Here it is more convenient to use the first and the third equations of the system for determining k_1 , k_2 and k_3 .

For $r_1 = -1$ we obtain

$$k_1^{(1)} + k_3^{(1)} = 0, \quad -k_2^{(1)} + k_3^{(1)} = 0$$

As solutions we can take, for example, the numbers $k_1^{(1)} = 1$, $k_2^{(1)} = -1$ and $k_3^{(1)} = -1$.

Consequently, the solution corresponding to $r_1 = 1$ is

$$x_1 = e^{-t}, \quad y_1 = -e^{-t}, \quad z_1 = -e^{-t}$$

For $r_2 = -2$ we have

$$2k_1^{(2)} + k_3^{(2)} = 0, \quad -k_1^{(2)} + 2k_3^{(2)} = 0$$

Here we can put $k_1^{(2)} = 1$, $k_2^{(2)} = -4$ and $k_3^{(2)} = -2$, then

$$x_2 = e^{-2t}, \quad y_2 = -4e^{-2t}, \quad z_2 = -2e^{-2t}$$

Finally, for $r_3 = 2$ we write

$$-2k_1^{(3)} + k_3^{(3)} = 0, \quad -k_1^{(3)} - 2k_3^{(3)} = 0$$

Taking $k_1^{(3)} = 1$, $k_2^{(3)} = -4$ and $k_3^{(3)} = 2$ we receive

$$x_3 = e^{2t}, \quad y_3 = -4e^{2t}, \quad z_3 = 2e^{2t}$$

Now, the general solution is written in the form

$$\begin{aligned} x &= C_1 e^{-t} + C_2 e^{-2t} + C_3 e^{2t} \\ y &= -C_1 e^{-t} - 4C_2 e^{-2t} - 4C_3 e^{2t} \\ z &= -C_1 e^{-t} - 2C_2 e^{-2t} + 2C_3 e^{2t} \end{aligned}$$

Let the reader check that particular solutions we have found form a fundamental system of solutions.

If among the simple roots of the characteristic equation there are *complex* ones the solution is also obtained in the form of complex functions of a real variable. In this case to find a real solution one should resort to the following property completely analogous to the one proved on page 610: *if complex functions $x_1 + ix_2$, $y_1 + iy_2$, $z_1 + iz_2$ are solutions of system (*) with real coefficients their real parts (x_1 , y_1 , z_1) and imaginary parts (x_2 , y_2 , z_2) are themselves solutions of the system.* In this situation to conjugate complex roots of the characteristic equation there correspond the same real solutions.

For the sake of brevity we shall demonstrate this case by the example of a system of two equations.

(2) Consider the system

$$\begin{aligned} \dot{x} &= -7x + y \\ \dot{y} &= -2x - 5y \end{aligned}$$

Its characteristic equation is

$$\begin{vmatrix} -7-r & 1 \\ -2 & -5-r \end{vmatrix} = r^2 + 12r + 37 = 0$$

whence $r_{1,2} = -6 \pm i$. Here the system of equations for determining k_1 and k_2 reduces to one equation with complex coefficients;

for the root $r_1 = -6 + i$ this equation is written as

$$(-1 - i)k_1 + k_2 = 0$$

(Let the reader check once again that the second equation of this algebraic system is a consequence of the first one.)

We can put $k_1 = 1$ and $k_2 = 1 + i$; then the complex expression of the solution of the given system of differential equations is

$$\begin{aligned} x &= e^{(-6+i)t} = e^{-6t} \cos t + ie^{-6t} \sin t \\ y &= (1+i)e^{(-6+i)t} = e^{-6t} (\cos t - \sin t) + ie^{-6t} (\cos t + \sin t) \end{aligned}$$

On separating the real and the imaginary parts we obtain two real solutions:

$$\begin{aligned} x_1 &= e^{-6t} \cos t, & y_1 &= e^{-6t} (\cos t - \sin t) \\ x_2 &= e^{-6t} \sin t, & y_2 &= e^{-6t} (\cos t + \sin t) \end{aligned}$$

As has been mentioned, the same operations on the conjugate complex root $r_2 = -6 - i$ lead to the same real solutions.

It is readily verified that the solutions we have found constitute a fundamental system, and therefore the general solution to the system is

$$\begin{aligned} x &= e^{-6t} (C_1 \cos t + C_2 \sin t) \\ y &= e^{-6t} [C_1 (\cos t - \sin t) + C_2 (\cos t + \sin t)] \end{aligned}$$

A particular solution of a system of nonhomogeneous equations can be found by reducing the system to one equation of a higher order and applying the techniques elaborated in Sec. 174 but we shall not dwell on this question in more detail.

Systems of linear differential equations with constant coefficients can also be solved by means of the methods of the operational calculus (mentioned at the end of Sec. 174); they are particularly convenient when it is required to determine a particular solution satisfying given initial conditions.

In conclusion we consider an electrotechnical problem leading to a system of linear equations with constant coefficients.

Let there be two circuits, their resistances, inductances and capacitances being, respectively, R_1 , L_1 , C_1 and R_2 , L_2 , C_2 . Suppose that there is an electromagnetic coupling between the circuits; this means that a change of the electric current in one of the circuits induces an electromotive force in the other and vice versa.

As is known from physics, under certain conditions, the induced electromotive force in the first circuit is equal to $-M \frac{di_2}{dt}$ and in the second circuit to $-M \frac{di_1}{dt}$ where M is a constant coefficient of mutual inductance and $i_1 = i_1(t)$ and $i_2 = i_2(t)$ are, respectively,

the current intensities in the first and in the second circuit. The coefficient M is positive if the directions of the electromotive forces of mutual inductance and self-inductance coincide and negative if otherwise.

Furthermore, suppose that there are no external sources of electric power in the circuits. Then the current flows in the circuits are described by the following system of differential equations (see Sec. 175, IV):

$$L_1 i_1' + R_1 i_1 + \frac{1}{C_1} \int i_1 dt + M i_2' = 0$$

$$L_2 i_2' + R_2 i_2 + \frac{1}{C_2} \int i_2 dt + M i_1' = 0$$

On differentiating with respect to t we receive

$$L_1 i_1'' + R_1 i_1' + \frac{1}{C_1} i_1 + M i_2'' = 0$$

$$L_2 i_2'' + R_2 i_2' + \frac{1}{C_2} i_2 + M i_1'' = 0$$

which is a system of two differential equations of the second order. It can be solved without the reduction to the normal form. Putting $i_1 = k_1 e^{rt}$ and $i_2 = k_2 e^{rt}$ we directly arrive at a system of homogeneous linear algebraic equations in k_1 and k_2 . Nontrivial (nonzero) solutions to the system exist if and only if the determinant of the system is equal to zero:

$$\begin{vmatrix} L_1 r^2 + R_1 r + \frac{1}{C_1} & M r^2 \\ M r^2 & L_2 r^2 + R_2 r + \frac{1}{C_2} \end{vmatrix} = 0$$

that is

$$(L_1 L_2 - M^2) r^4 + (L_1 R_2 + L_2 R_1) r^3 + \left(\frac{L_1}{C_2} + \frac{L_2}{C_1} + R_1 R_2 \right) r^2 + \left(\frac{R_1}{C_2} + \frac{R_2}{C_1} \right) r + \frac{1}{C_1 C_2} = 0$$

In many practical cases it turns out that the latter equation possesses two pairs of conjugate complex roots, which means that the oscillations of the currents in the circuits are superpositions of two harmonic vibrations. A thorough analysis of all the possible cases can be found in courses of electrical engineering. In conclusion we note that if the capacitances of the circuits are neglected the current flow is described by a system of equations of the first order and there are no vibration processes.

180*. Multiple Roots of the Characteristic Equation. If the characteristic equation possesses multiple roots the structure of the solution of the system becomes substantially more complicated.

In the case of a single higher-order differential equation we can immediately write down a fundamental system of solutions provided that the multiplicities of the roots of the characteristic equation are known (see Sec. 174) while in the case of a system the situation is more complicated. Without going into particulars we shall illustrate what has been said by two examples.

(1) Let us solve the system

$$\begin{aligned}\dot{x} &= 2x + y - z \\ \dot{y} &= x + 2y - z \\ \dot{z} &= 3x + 3y - 2z\end{aligned}$$

Its characteristic equation

$$\begin{vmatrix} 2-r & 1 & -1 \\ 1 & 2-r & -1 \\ 3 & 3 & -2-r \end{vmatrix} = -r(r^2 - 2r + 1) = 0$$

has the roots $r_1 = 0$ and $r_{2,3} = 1$. For the simple root $r_1 = 0$ we can put $k_1^{(1)} = k_2^{(1)} = 1$, $k_3^{(1)} = 3$. The corresponding solution can be taken as three constants $x_1 = 1$, $y_1 = 1$ and $z_1 = 3$ which of course can be multiplied by an arbitrary constant C_1 .

For the multiple root $r = 1$ the system of equations

$$k_1 + k_2 - k_3 = 0, \quad k_1 + k_2 - k_3 = 0, \quad 3k_1 + 3k_2 - 3k_3 = 0$$

for k_1, k_2, k_3 reduces to a single equation $k_1 + k_2 - k_3 = 0$ (in contrast to the case of the simple root when we obtain two independent equations). Here we can choose arbitrarily the values of two quantities; for instance, we can put $k_1 = C_2$ and $k_2 = C_3$. This results in $k_3 = C_2 + C_3$, and hence we arrive at a family of particular solutions dependent on two arbitrary constants:

$$x = C_2 e^t, \quad y = C_3 e^t, \quad z = (C_2 + C_3) e^t$$

It is clearly seen that it is a linear combination of the two particular solutions $x_2 = e^t, y_2 = 0, z_2 = e^t$ and $x_3 = 0, y_3 = e^t, z_3 = e^t$.

Combining the three solutions obtained we receive a system of solutions which is fundamental since the determinant

$$\begin{vmatrix} 1 & 1 & 3 \\ e^t & 0 & e^t \\ 0 & e^t & e^t \end{vmatrix} = e^{2t}$$

is nonzero. Finally, the general solution of the system is

$$x = C_1 + C_2 e^t, \quad y = C_1 + C_3 e^t, \quad z = 3C_1 + (C_2 + C_3) e^t$$

(2) Consider the system

$$\begin{aligned}\dot{x} &= x - y \\ \dot{y} &= 4x - y + 2z \\ \dot{z} &= 2y + z\end{aligned}$$

The roots of the characteristic equation

$$\begin{vmatrix} 1-r & -1 & 0 \\ 4 & -1-r & 2 \\ 0 & 2 & 1-r \end{vmatrix} = -(1-r)^2(1+r) = 0$$

are $r_1 = -1$ and $r_{2,3} = 1$. To the first simple root there corresponds the solution $x_1 = e^{-t}$, $y_1 = 2e^{-t}$, $z_1 = -2e^{-t}$.

Let us write down the system of equations for k_1 , k_2 and k_3 corresponding to the multiple root $r_{2,3} = 1$:

$$-k_2 = 0, \quad 4k_1 - 2k_2 + 2k_3 = 0, \quad 2k_2 = 0$$

We see that in this case it reduces to two equations, which give us only one particular solution; for instance, we can take $x_2 = e^t$, $y_2 = 0$ and $z_2 = -2e^t$. A third solution of the same structure which is not a linear combination of the two former solutions does not exist. It turns out that in this case it is possible to find a solution of the form

$$x = (\alpha_1 + \beta_1 t) e^t, \quad y = (\alpha_2 + \beta_2 t) e^t, \quad z = (\alpha_3 + \beta_3 t) e^t$$

The constants α_i , β_i ($i = 1, 2, 3$) are found with the aid of the method of undetermined coefficients. Indeed, on differentiating x , y and z with respect to t , substituting the results into the equations and cancelling by e^t we get the identities

$$\begin{aligned}\alpha_1 + \beta_1 t + \beta_1 &= \alpha_1 + \beta_1 t - \alpha_2 - \beta_2 t \\ \alpha_2 + \beta_2 t + \beta_2 &= 4(\alpha_1 + \beta_1 t) - \alpha_2 - \beta_2 t + 2(\alpha_3 + \beta_3 t) \\ \alpha_3 + \beta_3 t + \beta_3 &= 2(\alpha_2 + \beta_2 t) + \alpha_3 + \beta_3 t\end{aligned}$$

Now it only remains to equate in each identity the constant terms (this results in the equations written below on the left) and the coefficients in t (equations on the right):

$$\begin{array}{ll}(1) \beta_1 + \alpha_2 = 0 & (4) \beta_2 = 0 \\ (2) -4\alpha_1 + 2\alpha_2 + \beta_2 - 2\alpha_3 = 0 & (5) -4\beta_1 + 2\beta_2 - 2\beta_3 = 0 \\ (3) \beta_3 - 2\alpha_2 = 0 & (6) -2\beta_2 = 0\end{array}$$

Here the sixth equation is equivalent to the fourth one and the fifth is a consequence of the first, third and fourth equations. Thus, we arrive at the remaining four independent equations containing six unknowns. Setting the values of two of the unknowns, for instance $\alpha_1 = C_2$ and $\alpha_2 = C_3$ (where C_2 and C_3 are

of linear differential equations and their solutions in matrix form which is widely used in modern engineering literature.

Let there be given a homogeneous linear system of differential equations in normal form

$$\begin{aligned}\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\vdots \\ \dot{x}_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n\end{aligned}\quad (*)$$

The coefficients a_{ij} of the system may depend (continuously) on t or, in a particular case, be constant. Let us write down the so-called *coefficient matrix* of the system

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

and also the *column matrix* $X(t)$ whose elements are the functions $x_i(t)$ ($i=1, 2, \dots, n$) and the derivative $\dot{X}(t)$:

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad \dot{X}(t) = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}$$

The column matrix $X(t)$ is also termed a *vector function* of the scalar argument t , the functions $x_1(t)$, $x_2(t)$, $\dots$, $x_n(t)$ being its *coordinates*. In what follows we shall simply speak of the column matrix $X(t)$ as a vector and sometimes briefly denote it as X .

The product of the matrix A by the vector X is understood as a new vector specified by the relation

$$AX = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{pmatrix}$$

The elements of this vector written on the right exactly coincide with the right-hand sides of the system of equations (*) which therefore can be briefly written as one matrix differential equation:

$$\dot{X} = AX \quad (**)$$

Particular solutions of such an equation are vector functions (or, briefly, vectors) $X_i(t)$ with coordinates $x_{1i}(t)$, $x_{2i}(t)$, $\dots$, $x_{ni}(t)$.

Equation (**) apparently always possesses the zero (trivial) solution $\mathbf{X}_i \equiv 0$.

The rules of operations on matrices readily imply that any linear combination of particular solutions

$$C_1 \mathbf{X}_1(t) + C_2 \mathbf{X}_2(t) + \dots + C_m \mathbf{X}_m(t)$$

where all C_i 's are arbitrary constants is also a solution.

Let us state the definition of linear independence of a system of vector functions. *A system of vector functions is said to be linearly independent if the identity*¹

$$\lambda_1 \mathbf{X}_1(t) + \lambda_2 \mathbf{X}_2(t) + \dots + \lambda_m \mathbf{X}_m(t) \equiv 0 \quad (***)$$

is only possible for constant λ_i , $i = 1, 2, \dots, m$, when $\lambda_1 = \lambda_2 = \dots = \lambda_m = 0$. If otherwise, that is if this identity can hold when not all λ_i 's are equal to zero, the system is spoken of as *linearly dependent*.

In other words, a system of vector functions is linearly independent if none of the functions is representable as a linear combination of the others and is linearly dependent if otherwise. We also note that one vector identity (***) is equivalent to the n scalar identities

$$\begin{aligned} \lambda_1 x_{11}(t) + \lambda_2 x_{12}(t) + \dots + \lambda_m x_{1m}(t) &\equiv 0 \\ \lambda_1 x_{21}(t) + \lambda_2 x_{22}(t) + \dots + \lambda_m x_{2m}(t) &\equiv 0 \\ &\vdots \\ \lambda_1 x_{n1}(t) + \lambda_2 x_{n2}(t) + \dots + \lambda_m x_{nm}(t) &\equiv 0 \end{aligned}$$

containing the coordinates of the vectors. It is of course assumed that all the functions $x_{ij}(t)$ are defined and continuous in the same interval of variation of t .

Now let vector functions $\mathbf{X}_1(t)$, $\mathbf{X}_2(t)$, ..., $\mathbf{X}_n(t)$ be solutions to the matrix differential equation (**) or, which is the same, to system (*). For these solutions to be linearly independent it is necessary and sufficient that the determinant

$$W = \begin{vmatrix} x_{11}(t) & x_{12}(t) & \dots & x_{1n}(t) \\ x_{21}(t) & x_{22}(t) & \dots & x_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1}(t) & x_{n2}(t) & \dots & x_{nn}(t) \end{vmatrix}$$

be unequal to zero for all the values of t in which the coefficients $a_{ij}(t)$ of the system are continuous. In particular, if we deal with a system of n first-order linear differential equations, the condition should hold for the Wronskian theorem.

ing to the
are con-
h constant
we

¹ In this
of system (

)'s are any vector f

1)

Every linearly independent system of n solutions is called a *fundamental system*, and the determinant W is the *Wronskian* of that system. If a fundamental system of solutions $X_1(t)$, $X_2(t)$, ..., $X_n(t)$ is known the general solution is a linear combination of the constituent functions of the fundamental system:

$$X(t) = C_1 X_1(t) + C_2 X_2(t) + \dots + C_n X_n(t)$$

Let the reader write down the general solution in coordinate form involving n equalities containing the functions $x_{ij}(t)$.

As was indicated in Sec. 179, particular solutions of system (*) with constant coefficients are looked for in the form $x_1 = k_1 e^{rt}$, $x_2 = k_2 e^{rt}$, ..., $x_n = k_n e^{rt}$. In this case the vector functions $X(t)$ and $\dot{X}(t)$ are expressed as

$$X(t) = \begin{pmatrix} k_1 e^{rt} \\ k_2 e^{rt} \\ \vdots \\ k_n e^{rt} \end{pmatrix} = e^{rt} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}, \quad \dot{X}(t) = r e^{rt} \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix}$$

Let us denote the number vector with coordinates $k_1, k_2, \dots, k_n$ as K and substitute the above expressions of X and $\dot{X}$ into the matrix differential equation (**):

$$A e^{rt} K = r e^{rt} K$$

On cancelling the scalar factor e^{rt} we obtain the matrix algebraic equation

$$AK = rK \quad (****)$$

Now it is necessary to find a nonzero vector K (since if K is a zero vector we obtain the trivial solution we are not interested in) such that the multiplication of the matrix A by this vector yields a vector rK collinear with K . In the matrix calculus the numbers r for which there exist nonzero solutions of equation (****) are termed *eigenvalues* of the matrix A and the vectors K themselves are called its *eigenvectors*.

Note that if K is an eigenvector corresponding to the eigenvalue r then any vector λK collinear with it is also an eigenvector belonging to the same eigenvalue. For equation (****) implies that $A\lambda K = \lambda AK = \lambda rK = r\lambda K$.

In terms of the solutions of a system of differential equations this means that the multiplication of any solution by a scalar factor again results in a solution.

On writing the matrix equation for k_i in the coordinate form and transposing all the members to the left we obtain the homo-

10. Describe Euler's graphical method for constructing an integral curve of a differential equation of the first order.

11. Describe Euler's method of approximate numerical integration of an equation of the first order.

12*. What is a singular point of a differential equation of the first order? Describe various types of singular points and give illustrative examples.

13*. What solution of a differential equation is said to be singular? How is it found?

14. State the definition of a differential equation of the second order.

15. Explain the geometrical meaning of initial conditions for a differential equation of the second order.

16. State the theorem on the existence and uniqueness of the solution of an equation of the second order.

17. How is a second-order equation $y'' = f(x, y, y')$ reduced to an equation of the first order when its right-hand side does not contain (1) y and y' or (2) y or (3) x ?

18. Formulate the definition of a differential equation of the n th order and of its general solution. How are initial conditions set for an equation of the n th order?

19. What is a linear differential equation of the second order?

20. State and prove the theorem on the structure of the general solution of a homogeneous linear differential equation.

21. State and prove the theorem on the structure of the general solution of a nonhomogeneous linear equation.

22. Describe how a homogeneous linear equation of the second order with constant coefficients is solved. What is the characteristic equation? How is it formed?

23. Describe the structure of the general solution of a homogeneous linear differential equation of the second order with constant coefficients in the case of real and distinct roots of the characteristic equation and in the case of coinciding roots.

24. Prove that if a complex function is a solution to a linear differential equation with real coefficients its real and imaginary parts are also solutions.

25. What is the form of the solution in the case of complex roots of the characteristic equation?

26. What is the rule for determining a particular solution of a nonhomogeneous equation whose right-hand side is

$$f(x) = e^{px} [P_1(x) \cos qx + P_2(x) \sin qx]$$

where $P_1(x)$ and $P_2(x)$ are polynomials? Give examples.

27*. What is the idea of the method of complex amplitudes?

28. How can the solution of a nonhomogeneous equation be found when the right-hand side of the equation is a sum of several functions?

10. Describe Euler's graphical method for constructing an integral curve of a differential equation of the first order.

11. Describe Euler's method of approximate numerical integration of an equation of the first order.

12*. What is a singular point of a differential equation of the first order? Describe various types of singular points and give illustrative examples.

13*. What solution of a differential equation is said to be singular? How is it found?

14. State the definition of a differential equation of the second order.

15. Explain the geometrical meaning of initial conditions for a differential equation of the second order.

16. State the theorem on the existence and uniqueness of the solution of an equation of the second order.

17. How is a second-order equation $y'' = f(x, y, y')$ reduced to an equation of the first order when its right-hand side does not contain (1) y and y' or (2) y or (3) x ?

18. Formulate the definition of a differential equation of the n th order and of its general solution. How are initial conditions set for an equation of the n th order?

19. What is a linear differential equation of the second order?

20. State and prove the theorem on the structure of the general solution of a homogeneous linear differential equation.

21. State and prove the theorem on the structure of the general solution of a nonhomogeneous linear equation.

22. Describe how a homogeneous linear equation of the second order with constant coefficients is solved. What is the characteristic equation? How is it formed?

23. Describe the structure of the general solution of a homogeneous linear differential equation of the second order with constant coefficients in the case of real and distinct roots of the characteristic equation and in the case of coinciding roots.

24. Prove that if a complex function is a solution to a linear differential equation with real coefficients its real and imaginary parts are also solutions.

25. What is the form of the solution in the case of complex roots of the characteristic equation?

26. What is the rule for determining a particular solution of a nonhomogeneous equation whose right-hand side is

$$f(x) = e^{px} [P_1(x) \cos qx + P_2(x) \sin qx]$$

where $P_1(x)$ and $P_2(x)$ are polynomials? Give examples.

27*. What is the idea of the method of complex amplitudes?

28. How can the solution of a nonhomogeneous equation be found when the right-hand side of the equation is a sum of several functions?

29. Describe Lagrange's method of variation of arbitrary constants.

30. State the definition of a linearly independent system of functions. When is a system of functions said to be linearly dependent?

31. What condition is necessary and sufficient for a family of particular solutions of a linear differential equation of the n th order to be linearly independent?

32. Prove the theorem on the structure of the general solution for a homogeneous linear differential equation of the n th order and for a nonhomogeneous one.

33. How can the general solution of a homogeneous equation of the n th order with constant coefficients be formed depending on the roots of the characteristic equation?

34. What is a system of differential equations? Give the definition of a solution of such a system.

35. What is the normal form of a system of differential equations?

36. Describe the methods for reducing a differential equation of a higher order to a normal system of equations. How can the reverse transformation be performed? Is it always possible?

37. What is the structure of the general solution of a normal system of differential equations? State the existence and uniqueness theorem for the solution of such a system.

38*. Describe the geometrical interpretation of the solutions of a system of two differential equations as trajectories in the phase plane.

39*. Describe the properties of the solutions of an autonomous system of differential equations.

40. Enumerate the properties of the solutions of a system of linear differential equations.

41. Describe the method of solving a system of linear differential equations with constant coefficients with the aid of the characteristic equation. Consider the case of simple (real or complex) characteristic roots.

42*. What can be said about the structure of the solution in the case of multiple roots of the characteristic equation?

43*. Describe the matrix form of a system of linear differential equations. State the definition of a linearly independent system of vector functions and describe the structure of the general solution of the system.

44*. Establish the relationship between the problem of solving a homogeneous linear system of differential equations with constant coefficients and the problem of finding the eigenvalues and the eigenvectors of the coefficient matrix of the system.

Chapter XI

SERIES

§ 1. Numerical Series¹

182. **Series. Definition of the Sum of a Series.** Consider an infinite sequence of numbers $u_1, u_2, \dots, u_n, \dots$

Definition. An expression

$$u_1 + u_2 + \dots + u_n + \dots$$

is called a *(numerical) series*, and the numbers $u_1, u_2, \dots$ are called the *terms of the series*.

A series is briefly written as

$$\sum_{n=1}^{\infty} u_n$$

and u_n is called the *n*th term or the *general term* (*general element*) of the series.

A series is completely specified if there is a rule according to which for any number n the corresponding term of the series can be written: $u_n = f(n)$.

If a formula $u_n = f(n)$ is given we can immediately write any term of the series. For instance, if $u_n = \frac{1}{2^n}$, the series has the form

$$\frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^n} + \dots$$

If $u_n = \frac{1}{n!}$ the series is written as

$$1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} + \dots$$

A series is sometimes specified with the aid of a *recurrence formula* connecting its *n*th term with preceding ones. In this case

¹ The reader is advised to reread Secs. 26 and 31 before proceeding to study § 1 of this chapter.

we are given the first several terms of the series and a formula by which the other terms are found. For instance, let $u_1 = 1$ and $u_2 = \frac{1}{2}$, the recurrence formula being $u_n = \frac{1}{2}u_{n-1} + \frac{1}{3}u_{n-2}$. Using this formula we successively find

$$u_3 = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1 = \frac{7}{12}, \quad u_4 = \frac{1}{2} \cdot \frac{7}{12} + \frac{1}{3} \cdot \frac{1}{2} = \frac{11}{24}, \text{ etc.}$$

and thus obtain the series

$$1 + \frac{1}{2} + \frac{7}{12} + \frac{11}{24} + \dots$$

Let us consider a series

$$u_1 + u_2 + \dots + u_n + \dots$$

and designate the sum of its first n terms by s_n :

$$s_n = u_1 + u_2 + \dots + u_n$$

This sum is called the n th *partial sum of the series*.

Making n take on the values 1, 2, ... we obtain the sequence of partial sums of the series:

$$\begin{aligned} s_1 &= u_1 \\ s_2 &= u_1 + u_2 \\ &\dots \dots \dots \\ s_n &= u_1 + u_2 + \dots + u_n \\ &\dots \dots \dots \end{aligned}$$

As n increases indefinitely a greater and still greater number of terms of the series is involved in the sum s_n . Thus we naturally come to the following definition:

Definition. If the sequence of partial sums of the given series has a definite limit as $n \rightarrow \infty$, i.e.

$$\lim_{n \rightarrow \infty} s_n = s$$

the series is said to be *convergent* and the number s is called the *sum of the series*.

In this case we write

$$s = u_1 + u_2 + \dots + u_n + \dots$$

If the sequence s_n does not tend to any limit, the series is said to be *divergent*.

It should be stressed that a series can be divergent in the following two cases: (1) when the sequence s_n tends to infinity, (2) when the sequence s_n tends to a finite but not infinite limit.

As a simple example, let us consider an infinite geometrical series:

$$a + aq + aq^2 + \dots + aq^{n-1} + \dots \quad (a \neq 0) \quad (*)$$

The sum of the first n terms of the series is

$$s_n = a \frac{q^n - 1}{q - 1}$$

If $|q| < 1$ then $\lim_{n \rightarrow \infty} q^n = 0$, and therefore

$$\lim_{n \rightarrow \infty} s_n = a \lim_{n \rightarrow \infty} \frac{q^n - 1}{q - 1} = \frac{a}{1 - q}$$

Consequently, an infinite geometric series with $|q| < 1$ is convergent and its sum is given by

$$s = \frac{a}{1 - q}$$

If $|q| > 1$ then $\lim_{n \rightarrow \infty} q^n = \infty$, and hence, $\lim_{n \rightarrow \infty} s_n = \infty$, that is, the series is divergent.

Now, let $q = 1$. The series

$$a + a + a + \dots + a + \dots \quad (a \neq 0)$$

has the n th partial sum $s_n = na$ which tends to infinity together with n : $\lim_{n \rightarrow \infty} s_n = \infty$.

If $q = -1$ we obtain the series

$$a - a + a - a + \dots$$

Its partial sums take on the values $s_1 = a$, $s_2 = 0$, $s_3 = a$, $s_4 = 0$, etc., that is, s_n does not tend to any limit. According to the definition, this series is divergent.

Thus, an infinite geometric series is convergent for $|q| < 1$ and divergent for $|q| \geq 1$.

In this example we find out whether the series is convergent or divergent by using the definition of convergence and the known formula for the n th partial sum.

However, in many cases this approach proves inapplicable because it is very difficult and sometimes even impossible to derive a formula for s_n and to determine the limit of s_n as $n \rightarrow \infty$. Therefore the convergence of a series is usually investigated without finding its sum by using certain *sufficient conditions (tests) for convergence* which will be established later. Before proceeding to these sufficient conditions we shall explain why it is so important to know whether the series is convergent or divergent even when we cannot find its sum.

Let us consider a convergent series

$$s = u_1 + u_2 + \dots + u_n + \dots$$

The difference between the sum of the series and its n th partial sum is called the *remainder after n terms* (or the *n th remainder*). The n th remainder of a convergent series $u_1 + u_2 + \dots + u_n + \dots$ is the sum of the infinite series $u_{n+1} + u_{n+2} + \dots$ ¹; we designate it by r_n . We have

$$r_n = s - s_n = u_{n+1} + u_{n+2} + \dots$$

The original series is assumed to be convergent, that is $\lim_{n \rightarrow \infty} s_n = s$. Consequently, the absolute value of the remainder

$$|r_n| = |s - s_n|$$

becomes arbitrarily small when the number n is sufficiently large. Thus, it is always possible to compute an *approximation* to the sum of a convergent series by taking a sufficiently large number of its first terms. However, in the general case it is extremely difficult to determine the error of such an approximation. Later on we shall present some examples showing how in certain cases the error can be *estimated* and how the number of terms of the series required for obtaining its sum with a given accuracy can be determined. It should be stressed once again that a *divergent* series has no sum.

Let us enumerate the simplest properties of convergent series.

1. If a series

$$u_1 + u_2 + \dots + u_n + \dots$$

is convergent, its sum being s , the series

$$\lambda u_1 + \lambda u_2 + \dots + \lambda u_n + \dots$$

composed of the products of the terms of the former series by an arbitrary number λ is also convergent and its sum is equal to λs .

To prove this assertion we designate the n th partial sum of the former series by s_n and that of the latter by σ_n . Then

$$\sigma_n = \lambda u_1 + \lambda u_2 + \dots + \lambda u_n = \lambda s_n$$

Consequently,

$$\lim_{n \rightarrow \infty} \sigma_n = \lim_{n \rightarrow \infty} (\lambda s_n) = \lambda \lim_{n \rightarrow \infty} s_n = \lambda s$$

The property has been proved.

¹ Its convergence will be proved later (see Property 3). It should be noted that the term "remainder" is sometimes used both for the series $u_{n+1} + u_{n+2} + \dots$ and for its sum. It would be more precise to call $u_{n+1} + u_{n+2} + \dots$ "the remainder series" and the sum of this series "the remainder" (or "the remainder term").

2. If

$$s' = u_1 + u_2 + \dots + u_n + \dots$$

and

$$s'' = v_1 + v_2 + \dots + v_n + \dots$$

are two convergent series the series formed of the sums of the corresponding terms of the given series is also convergent and its sum is equal to $s' + s''$:

$$(u_1 + v_1) + (u_2 + v_2) + \dots + (u_n + v_n) + \dots = s' + s''$$

The proof of this property is also very simple; it is based on the theorem on the limit of a sum. Let s'_n and s''_n be the n th partial sums of the two given series, and let σ_n be the n th partial sum of the resultant series. Then

$$\sigma_n = (u_1 + v_1) + \dots + (u_n + v_n) = s'_n + s''_n$$

whence

$$\lim_{n \rightarrow \infty} \sigma_n = \lim_{n \rightarrow \infty} (s'_n + s''_n) = \lim_{n \rightarrow \infty} s'_n + \lim_{n \rightarrow \infty} s''_n = s' + s''$$

which is what we wanted to prove.

On multiplying the terms of the second series by -1 (by Property 1, the sum of the new series is then equal to $-s''$) and adding them to the corresponding terms of the first series we see that

$$(u_1 - v_1) + (u_2 - v_2) + \dots + (u_n - v_n) + \dots = s' - s''$$

3. If a series is convergent the series obtained by adding or deleting a finite number of terms is also convergent.

Suppose that we have deleted a finite number of terms of a convergent series

$$u_1 + u_2 + \dots + u_n + \dots$$

For instance let these terms be u_2 , u_8 , u_{10} and u_{15} . To retain the indices of the remaining terms let us agree that the deleted terms have been replaced by zeros. Then, for $n > 15$ the partial sums of both series will differ from each other in the constant summand $u_2 + u_8 + u_{10} + u_{15}$. Therefore, if one of these partial sums has a limit the other sum also tends to a limit, their difference being equal to the same constant summand.¹

Analogous considerations apply to the case when some new terms are added to the given series.

A consequence of this property is that *if a series is convergent its any remainder series is also convergent and vice versa.*

¹ If the sum of the deleted terms is equal to zero the sums of both series coincide.

183. **Necessary Condition for Convergence of a Series.** Harmonic Series. As was mentioned, the problem of prime importance in the investigation of a series is to find out whether it is convergent or divergent. Therefore, our aim is to establish some conditions (tests) which enable us to judge upon the convergence or divergence of a series by inspecting its general term.

Necessary Condition for Convergence of a Series. If a series is convergent its n th term tends to zero as $n \rightarrow \infty$.

Proof. We have

$$s_n = u_1 + u_2 + \dots + u_{n-1} + u_n = s_{n-1} + u_n$$

If the series is convergent we have

$$\lim_{n \rightarrow \infty} s_{n-1} = s \quad \text{and} \quad \lim_{n \rightarrow \infty} s_n = s$$

Since

$$\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1}$$

we obtain

$$\lim_{n \rightarrow \infty} u_n = s - s = 0$$

The necessary condition we have proved implies the following *sufficient test for divergence of a series*:

If u_n does not tend to zero as $n \rightarrow \infty$ the series cannot be convergent, that is, it must be divergent.

As an example, let us take the series

$$\frac{1}{101} + \frac{2}{201} + \frac{3}{301} + \dots + \frac{n}{100n+1} + \dots$$

This series is divergent because its general term $u_n = \frac{n}{100n+1}$ tends to $\frac{1}{100}$ as $n \rightarrow \infty$ and hence does not tend to zero.

It should be stressed that the fact that the n th term of a series tends to zero as $n \rightarrow \infty$ is not sufficient for the convergence of the series.

To elucidate what has been said let us consider the so-called *harmonic series*:

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$$

Here we have $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} \frac{1}{n} = 0$ but the series is divergent. Indeed, as will be proved, in this case $\lim_{n \rightarrow \infty} s_n = +\infty$.

To prove the divergence we shall replace some terms of the series by smaller numbers and then show that the sum of the smaller summands tends to infinity. On grouping the terms of

harmonic series we can write it as

$$1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \left(\frac{1}{9} + \dots + \frac{1}{16}\right) + \dots$$

In each bracket we now replace all the summands by the last one which remains unchanged. This results in

$$1 + \frac{1}{2} + \underbrace{\left(\frac{1}{4} + \frac{1}{4}\right)}_2 + \underbrace{\left(\frac{1}{8} + \dots + \frac{1}{8}\right)}_4 + \underbrace{\left(\frac{1}{16} + \dots + \frac{1}{16}\right)}_8 + \dots$$

In each bracket we have written the number of summands it contains. It is obvious that this operation decreases the number of terms in each bracket and reduces it to $\frac{1}{2}$. Since we can use an arbitrary number of such brackets their sum tends to infinity.

It follows that the n th partial sum of the harmonic series increases indefinitely as $n \rightarrow \infty$ and, consequently, the series is divergent.

It is still easier to prove the divergence of the series $\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \dots + \frac{1}{\sqrt{n}} + \dots$ whose n th term also tends to zero. Here we have

$$= 1 + \frac{1}{\sqrt{2}} + \dots + \frac{1}{\sqrt{n}} > \frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n}} + \dots + \frac{1}{\sqrt{n}} = \frac{n}{\sqrt{n}} = \sqrt{n}$$

Consequently, $\lim_{n \rightarrow \infty} s_n = \infty$.

4. Positive Series. Sufficient Conditions for Convergence.

Comparison tests. Let us consider the so-called *positive series* whose all terms are *positive*:

$$u_1 + u_2 + \dots + u_n + \dots \quad (u_n > 0)$$

The investigation of such series is facilitated by the following proposition.

Lemma. If the partial sums of a positive series are bounded above, that is

$$s_n < M, \quad M = \text{const}$$

the series is convergent¹.

Proof. Since all the terms of the series are positive, its n th partial sum increases together with the number n :

$$s_1 < s_2 < \dots < s_n < \dots$$

¹This condition does not imply the convergence of a series with arbitrary terms. For instance, as was shown, the series $a - a + a - a + \dots$ ($a > 0$) whose partial sums do not exceed the number a is divergent.

But if a monotone increasing sequence is bounded above by a constant number it has a limit (see Sec. 31) which does not exceed that number either. Hence,

$$\lim_{n \rightarrow \infty} s_n = s \leq M$$

Conversely,

If a positive series is convergent its partial sums are less than the sum of the series: $s_n < s$.

It should also be noted that if a positive series is divergent its partial sums tend to infinity: $\lim_{n \rightarrow \infty} s_n = +\infty$. In this case we

write $\sum_{n=1}^{\infty} u_n = +\infty$ and say that the series is *divergent to $+\infty$* .

The lemma we have proved directly implies the following comparison tests for positive series.

Comparison Tests. Consider two positive series

$$\sum_{k=1}^{\infty} u_k = u_1 + u_2 + \dots + u_k + \dots \quad (u_k > 0) \quad (1)$$

and

$$\sum_{k=1}^{\infty} v_k = v_1 + v_2 + \dots + v_k + \dots \quad (v_k > 0) \quad (2)$$

satisfying the condition that each term of series (1) does not exceed the corresponding term of series (2), i.e.

$$u_k \leq v_k \quad (k = 1, 2, \dots) \quad (*)$$

Then:

(1) If series (2) is convergent series (1) is also convergent.

(2) If series (1) is divergent series (2) is also divergent.

Proof. We shall begin with the first part of the proposition. Let us put

$$s_n = \sum_{k=1}^n u_k, \quad \sigma_n = \sum_{k=1}^n v_k$$

By the hypothesis, series (2) is convergent and therefore $\sigma_n < \sigma$ where σ is the sum of the second series. It follows from condition (*) that

$$s_n \leq \sigma_n < \sigma$$

which means that the partial sums of series (1) are bounded above. But then the lemma we have proved implies that series (1) is convergent.

To prove the second part of the proposition we make use of the fact that if series (1) is divergent its partial sums increase indefinitely:

Since $\sigma_n \geq s_n$ we also have $\sigma_n \rightarrow \infty$. Consequently, series (2) is also divergent, which is what we had to prove.

Comparison tests are also applicable when the terms of the series satisfy condition (*) not for all n but only beginning with a certain number $n = N$. This is implied by the third property of convergent series (see Sec. 182).

As an instance, let us prove that the series

$$1 + \frac{1}{\sqrt{2}} + \dots + \frac{1}{\sqrt{n}} + \dots$$

is divergent. To this end it is sufficient to compare the given series with the harmonic series

$$1 + \frac{1}{2} + \dots + \frac{1}{n} + \dots$$

Since $\frac{1}{\sqrt{n}} > \frac{1}{n}$ for $n > 1$ and the harmonic series is divergent the given series is also divergent. (On page 665 we demonstrated another proof of the divergence of this series.)

It is also easy to prove the convergence of the series

$$\frac{1}{2} + \frac{1}{2 \cdot 2^2} + \frac{1}{3 \cdot 2^3} + \dots + \frac{1}{n \cdot 2^n} + \dots$$

whose terms are smaller than the corresponding terms of the convergent geometric series

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^n} + \dots$$

In practice comparison tests can be conveniently used in the limiting form: if the limit of the ratio of u_n to v_n exists and is not equal to zero, i.e.

$$\lim_{n \rightarrow \infty} \frac{u_n}{v_n} = A > 0$$

series (1) and (2) are simultaneously convergent or divergent.

Let us prove this proposition. From the definition of a limit it follows that, given any $\varepsilon > 0$, there is a number N such that for all $n > N$ the inequality $\left| \frac{u_n}{v_n} - A \right| < \varepsilon$ holds or, equivalently,

$$A - \varepsilon < \frac{u_n}{v_n} < A + \varepsilon$$

where ε is assumed to be so small that $A - \varepsilon > 0$.

Let us suppose that series (2) is convergent; then, by Property 1 stated in Sec. 182, the series with general term $(A + \varepsilon)v_n$ is convergent and, by the comparison test, series (1) is also convergent because $u_n < (A + \varepsilon)v_n$ for all $n > N$.

But if a monotone increasing sequence is bounded above by a constant number it has a limit (see Sec. 31) which does not exceed that number either. Hence,

$$\lim_{n \rightarrow \infty} s_n = s \leq M$$

Conversely,

If a positive series is convergent its partial sums are less than the sum of the series: $s_n < s$.

It should also be noted that if a positive series is divergent its partial sums tend to infinity: $\lim_{n \rightarrow \infty} s_n = +\infty$. In this case we

write $\sum_{n=1}^{\infty} u_n = +\infty$ and say that the series is *divergent to $+\infty$* .

The lemma we have proved directly implies the following comparison tests for positive series.

Comparison Tests. Consider two positive series

$$\sum_{k=1}^{\infty} u_k = u_1 + u_2 + \dots + u_k + \dots \quad (u_k > 0) \quad (1)$$

and

$$\sum_{k=1}^{\infty} v_k = v_1 + v_2 + \dots + v_k + \dots \quad (v_k > 0) \quad (2)$$

satisfying the condition that each term of series (1) does not exceed the corresponding term of series (2), i.e.

$$u_k \leq v_k \quad (k = 1, 2, \dots) \quad (*)$$

Then:

(1) If series (2) is convergent series (1) is also convergent.

(2) If series (1) is divergent series (2) is also divergent.

Proof. We shall begin with the first part of the proposition.

Let us put

$$s_n = \sum_{k=1}^n u_k, \quad \sigma_n = \sum_{k=1}^n v_k$$

By the hypothesis, series (2) is convergent and therefore $\sigma_n < \sigma$ where σ is the sum of the second series. It follows from condition (*) that

$$s_n \leq \sigma_n < \sigma$$

which means that the partial sums of series (1) are bounded above. But then the lemma we have proved implies that series (1) is convergent.

To prove the second part of the proposition we make use of the fact that if series (1) is divergent its partial sums increase indefinitely:

$$s_n \rightarrow \infty$$

Since $\sigma_n \geq s_n$ we also have $\sigma_n \rightarrow \infty$. Consequently, series (2) is also divergent, which is what we had to prove.

Comparison tests are also applicable when the terms of the series satisfy condition (*) not for all n but only beginning with a certain number $n = N$. This is implied by the third property of convergent series (see Sec. 182).

As an instance, let us prove that the series

$$1 + \frac{1}{\sqrt{2}} + \dots + \frac{1}{\sqrt{n}} + \dots$$

is divergent. To this end it is sufficient to compare the given series with the harmonic series

$$1 + \frac{1}{2} + \dots + \frac{1}{n} + \dots$$

Since $\frac{1}{\sqrt{n}} > \frac{1}{n}$ for $n > 1$ and the harmonic series is divergent the given series is also divergent. (On page 665 we demonstrated another proof of the divergence of this series.)

It is also easy to prove the convergence of the series

$$\frac{1}{2} + \frac{1}{2 \cdot 2^2} + \frac{1}{3 \cdot 2^3} + \dots + \frac{1}{n \cdot 2^n} + \dots$$

whose terms are smaller than the corresponding terms of the convergent geometric series

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^n} + \dots$$

In practice comparison tests can be conveniently used in the limiting form: if the limit of the ratio of u_n to v_n exists and is not equal to zero, i.e.

$$\lim_{n \rightarrow \infty} \frac{u_n}{v_n} = A > 0$$

series (1) and (2) are simultaneously convergent or divergent.

Let us prove this proposition. From the definition of a limit it follows that, given any $\epsilon > 0$, there is a number N such that for all $n > N$ the inequality $\left| \frac{u_n}{v_n} - A \right| < \epsilon$ holds or, equivalently,

$$A - \epsilon < \frac{u_n}{v_n} < A + \epsilon$$

where ϵ is assumed to be so small that $A - \epsilon > 0$.

Let us suppose that series (2) is convergent; then, by Property 1 stated in Sec. 182, the series with general term $(A + \epsilon)v_n$ is convergent and, by the comparison test, series (1) is also convergent because $u_n < (A + \epsilon)v_n$ for all $n > N$.

Conversely, if series (2) is divergent, the inequality $u_n > (A - \epsilon)v_n$ implies that series (1) is divergent as well.

For instance, the series with the general term $u_n = \frac{2n}{3n^2 - 1}$ is divergent because the limit of the ratio of u_n to $\frac{1}{n}$ is equal to $\frac{2}{3}$.

A difficulty in applying the comparison tests is that for the given series it is necessary to construct another series with which the former can be compared. The harmonic series is often taken as a "standard" divergent series, whereas as a "standard" convergent series we usually take a convergent geometric series or the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots + \frac{1}{n^2} + \dots$$

The convergence of the latter series can be proved in the following way: its terms are combined into groups containing 1, 2, 4, 8, ... successive terms of the series, and in each group, beginning with the second, all the summands are replaced by the largest one, that is by the first term of the group. This leads to the obvious inequalities

$$\begin{aligned} \frac{1}{2^2} + \frac{1}{3^2} &< \frac{1}{2^2} + \frac{1}{2^2} = \frac{1}{2}, \quad \frac{1}{4^2} + \frac{1}{5^2} + \frac{1}{6^2} + \frac{1}{7^2} < 4 \cdot \frac{1}{4^2} = \frac{1}{4} = \frac{1}{2^2} \\ \frac{1}{8^2} + \dots + \frac{1}{15^2} &< 8 \cdot \frac{1}{8^2} = \frac{1}{8} = \frac{1}{2^3}, \text{ etc.} \end{aligned}$$

It follows that the partial sums of the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ are bounded (they do not exceed the sum of the terms of the geometric series $1 + \frac{1}{2} + \frac{1}{4} + \dots = 2$) and hence the series is convergent. As will be shown in Sec. 202, the sum of this series is equal to $\frac{\pi^2}{6} \approx 1.6449$.

As an instance, with the aid of the latter series we can easily prove the convergence of the series with the general term $u_n = \frac{n}{n^2 + 1}$. Indeed, the limit of the ratio of u_n to $\frac{1}{n^2}$ is equal to 1.

Now we proceed to consider a test which is in many cases applicable to proving the convergence of a series using solely the expression of its general term.

II. D'Alembert's¹ Test. Suppose we are given a positive series

$$u_1 + u_2 + \dots + u_n + \dots \quad (u_n > 0) \quad (*)$$

¹ D'Alembert, J. (1717-1783) a French mathematician and philosopher.

If, for $n \rightarrow \infty$, there exists a limit ρ of the ratio of two consecutive terms of the series, that is

$$\lim_{n \rightarrow \infty} \frac{u_{n+1}}{u_n} = \rho$$

then the series is convergent for $\rho < 1$ and divergent for $\rho > 1$.¹

For $\rho = 1$ D'Alembert's test is inapplicable (i.e. there are both convergent and divergent positive series with $\rho = 1$).

Proof. Let $\rho < 1$. By the definition of a limit, it is always possible to choose a number N such that for all $n \geq N$ there holds the inequality

$$\frac{u_{n+1}}{u_n} < \rho + \varepsilon = \rho_1$$

where $\varepsilon > 0$ is chosen sufficiently small so that $\rho_1 = \rho + \varepsilon$ is smaller than 1. Then

$$\frac{u_{N+1}}{u_N} < \rho_1, \quad \frac{u_{N+2}}{u_{N+1}} < \rho_1, \quad \frac{u_{N+3}}{u_{N+2}} < \rho_1, \quad \text{etc.}$$

We have

$$\begin{aligned} u_{N+1} &< \rho_1 u_N \\ u_{N+2} &< \rho_1 u_{N+1} < \rho_1^2 u_N \\ u_{N+3} &< \rho_1 u_{N+2} < \rho_1^3 u_N \\ &\dots \end{aligned}$$

It follows that the terms of the series

$$u_{N+1} + u_{N+2} + u_{N+3} + \dots$$

representing the remainder after the N th term of the given series are smaller than the corresponding terms of the convergent geometric series

$$\rho_1 u_N + \rho_1^2 u_N + \rho_1^3 u_N + \dots$$

(its common ratio ρ_1 is smaller than 1). Consequently, the remainder after the N th term is convergent, and therefore the given series itself is also convergent.

Now let $\rho > 1$ (in particular, there can be $\rho = \infty$). Then there is a number N such that for all $n > N$ the inequalities

$$\frac{u_{n+1}}{u_n} > 1, \quad \text{i. e. } u_{n+1} > u_n$$

hold. This means that every term of the series is greater than the preceding one and, since they are all positive, the necessary condition for convergence (requiring that the general term must tend to zero) is violated. Hence, the series is divergent.

¹ The series is also divergent if $\frac{u_{n+1}}{u_n} \rightarrow +\infty$ for $n \rightarrow \infty$; in this case we say that $\rho = \infty$.

Thus, if the divergence of series (*) is established with the aid of D'Alembert's test (i. e. $\rho > 1$) this means that u_n does not tend to zero as $n \rightarrow \infty$.

In the case when the limit ρ is equal to unity D'Alembert's test is inapplicable: there exist both convergent and divergent series for which $\rho = 1$.

Examples. (1) Let us consider the series $\sum_{n=1}^{\infty} \frac{n}{2^n}$. Here we have $u_n = \frac{n}{2^n}$ and $u_{n+1} = \frac{n+1}{2^{n+1}}$. Therefore $\rho = \lim_{n \rightarrow \infty} \frac{\frac{n+1}{2^{n+1}}}{\frac{n}{2^n}} = \frac{1}{2} \lim_{n \rightarrow \infty} \frac{n+1}{n} = \frac{1}{2}$.

Consequently, the series is convergent.

(2) Let us take the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ (called the p series). In this case $\frac{u_{n+1}}{u_n} = \frac{n^p}{(n+1)^p} = \left(\frac{n}{n+1}\right)^p \rightarrow 1$ for $n \rightarrow \infty$ irrespective of the exponent p . For $p=1$ we obtain the harmonic series which, as was proved, is divergent. For $p=2$ we obtain the convergent series $\sum_{n=1}^{\infty} \frac{1}{n^2}$. We see that when $\rho=1$ it is impossible to judge upon the convergence or divergence of this series by D'Alembert's test. In this case the series should be investigated with the aid of some other tests.

III. **Cauchy's root test.** Here we shall state without proof another sufficient test (known as *Cauchy's root test*) for convergence of positive series (*).

Suppose that the limit

$$\lim_{n \rightarrow \infty} \sqrt[n]{a_n} = l$$

(finite or infinite) exists. Then, if $l < 1$ the series is convergent and if $l > 1$ the series is divergent.

In the case $l=1$ this test does not allow us to judge upon the convergence of the series.

As an example, consider the series $\frac{1}{\ln 2} + \frac{1}{\ln^2 3} + \dots + \frac{1}{\ln^n(n+1)} + \dots$; it is convergent because $\lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{\ln^n(n+1)}} = \lim_{n \rightarrow \infty} \frac{1}{\ln(n+1)} = 0$.

It should be noted that there exist some more sophisticated sufficient tests which sometimes enable us to prove the convergence or divergence of a series when the tests considered here fail.

185. Cauchy's Integral Test. The reader must have already noticed a close analogy between the definitions of the convergence

of a series and of the convergence of an improper integral with infinite upper limit (see Sec. 97). The tests for convergence of positive series and for convergence of integrals with positive integrands also have very much in common. In this section we shall present a test enabling us to reduce, in some cases, the problem of convergence of a series to the problem of convergence of an improper integral, which may prove simpler.

Cauchy's Integral Test. Consider a series

$$u_1 + u_2 + \dots + u_n + \dots \quad (u_n > 0)$$

whose terms are the values of a continuous positive monotone increasing function $f(x)$ defined in the interval $[1, \infty]$ which are assumed for the integral values of the argument x :

$$u_1 = f(1), \quad u_2 = f(2), \quad \dots, \quad u_n = f(n), \quad \dots$$

Then the given series and the improper integral $\int_1^{\infty} f(x) dx$ are simultaneously convergent or divergent.

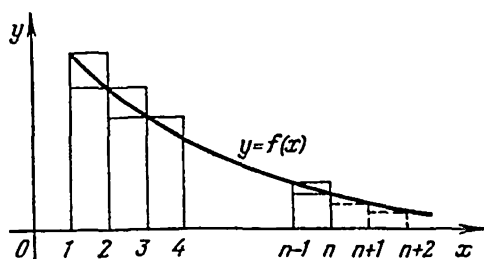


Fig. 246

Proof. Let us consider the curvilinear trapezoid bounded by the line $y=f(x)$, with the segment $[1, n]$ of the x -axis as base, where n is an arbitrary positive integer (Fig. 246). Its area is equal to the integral

$$I_n = \int_1^n f(x) dx$$

Let us construct the ordinates of the function $y=f(x)$ at the points $x=1, x=2, \dots, x=n-1, x=n$ and consider the two step figures shown in Fig. 246: one of them, contained within the trapezoid, has the area equal to $f(2) + f(3) + \dots + f(n) = s_n - u_1$ and the other (enveloping the trapezoid) has the area $f(1) + f(2) + \dots + f(n-1) = s_n - u_n$ where $s_n = u_1 + u_2 + \dots + u_n$. The area of the former figure is smaller than the area of the curvilinear trapezoid and

the area of the latter is greater than the area of the trapezoid

$$s_n - u_1 < I_n < s_n - u_n$$

This leads to the two inequalities

$$s_n < I_n + u_1 \quad (*)$$

and

$$s_n > I_n + u_n \quad (**)$$

The function $f(x)$ being positive, the integral I_n increases together with n ; there are two possible cases here:

(1) The improper integral is convergent, that is $I = \lim_{n \rightarrow \infty} I_n$ exists; then $I_n < I$, and inequality (*) indicates that

$$s_n < u_1 + I$$

for any n . Consequently, the partial sums s_n are bounded and, by the lemma in Sec. 184, the series is convergent.

(2) The integral is divergent; then $I_n \rightarrow \infty$ for $n \rightarrow \infty$, and we conclude, by inequality (**), that s_n also increases indefinitely, i.e. the series is divergent.

We shall demonstrate by the example of the p series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \dots + \frac{1}{n^p} + \dots$$

that the integral test can give us the required answer when D'Alembert's test is inapplicable. As was shown, for this series we have $\frac{u_{n+1}}{u_n} \rightarrow 1$; let us apply to it the integral test. It is obvious that as $f(x)$ we can take the function $\frac{1}{x^p}$. The integral

$$\int_1^{\infty} \frac{1}{x^p} dx = \left. \frac{x^{-p+1}}{-p+1} \right|_1^{\infty} = \frac{1}{1-p} \lim_{x \rightarrow \infty} (x^{-p+1} - 1)$$

is convergent if $p > 1$ because in this case $\lim_{x \rightarrow \infty} x^{-p+1} = 0$ and

$\int_1^{\infty} \frac{dx}{x^p} = \frac{1}{p-1}$ if $p < 1$ the integral is divergent; in the latter case $\lim_{x \rightarrow \infty} x^{-p+1} = \infty$ is also

divergent $\lim_{n \rightarrow \infty} n x^{-p+1} = \infty$. The series is divergent

For $p=1$ we obtain the harmonic series, which once again proves its divergence. For $p=2$ we obtain the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ whose convergence was already established.

The geometrical idea of the proof of Cauchy's integral test can be applied to estimate the error appearing in an approximation to the sum of a series whose convergence has been proved with the aid of this test. As was indicated on p. 662, to estimate the error we should find an estimation for the difference $s-s_n$, i.e. for the remainder

$$r_n = u_{n+1} + u_{n+2} + \dots$$

(Since we deal with positive series the difference $s-s_n$ is positive for any n .)

Fig. 246 indicates that the remainder is equal to the sum of the areas of the rectangles "infinitely receding to the right" (they are shown in dotted line); this sum is less than the area of the curvilinear figure (bounded by the line $y=f(x)$) enclosing these rectangles, and therefore

$$r_n < \int_n^{\infty} f(x) dx$$

Let us apply this formula to estimate the remainder of the p series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ for $p > 1$:

$$r_n < \int_n^{\infty} \frac{dx}{x^p} = \left. \frac{x^{-p+1}}{-p+1} \right|_n^{\infty} = \frac{1}{(p-1)n^{p-1}}$$

Hence, given an arbitrary number $\varepsilon > 0$, for the remainder to be not greater than ε , the number n can be put equal to the least integer satisfying the inequality

$$\frac{1}{(p-1)n^{p-1}} \leq \varepsilon$$

For instance, let $p=2$ and $\varepsilon=0.001$. Then the inequality $\frac{1}{n} \leq 0.001$ is fulfilled for $n=1000$. This estimation shows that, as we say, the series in question is *slowly convergent*: to find its sum with an error not exceeding 0.001 we should compute the sum of its first 1000 terms! Besides, to guarantee the accuracy of the result to the required number of correct decimal digits we should compute the terms of the series with a very high precision. Such calculations have of course no practical value.

the area of the latter is greater than the area of the trapezoid:

$$s_n - u_1 < I_n < s_n - u_n$$

This leads to the two inequalities

$$s_n < I_n + u_1 \quad (*)$$

and

$$s_n > I_n + u_n \quad (**)$$

The function $f(x)$ being positive, the integral I_n increases together with n ; there are two possible cases here:

(1) The improper integral is convergent, that is $I = \lim_{n \rightarrow \infty} I_n$ exists; then $I_n < I$, and inequality (*) indicates that

$$s_n < u_1 + I$$

for any n . Consequently, the partial sums s_n are bounded and, by the lemma in Sec. 184, the series is convergent.

(2) The integral is divergent; then $I_n \rightarrow \infty$ for $n \rightarrow \infty$, and we conclude, by inequality (**), that s_n also increases indefinitely, i.e. the series is divergent.

We shall demonstrate by the example of the p series

$$\sum_{n=1}^{\infty} \frac{1}{n^p} = 1 + \frac{1}{2^p} + \frac{1}{3^p} + \dots + \frac{1}{n^p} + \dots$$

that the integral test can give us the required answer when D'Alembert's test is inapplicable. As was shown, for this series we have $\frac{u_{n+1}}{u_n} \rightarrow 1$; let us apply to it the integral test. It is obvious that as $f(x)$ we can take the function $\frac{1}{x^p}$. The integral

$$\int_1^{\infty} \frac{1}{x^p} dx = \left. \frac{x^{-p+1}}{-p+1} \right|_1^{\infty} = \frac{1}{1-p} \lim_{x \rightarrow \infty} (x^{-p+1} - 1)$$

is convergent if $p > 1$ because in this case $\lim_{x \rightarrow \infty} x^{-p+1} = 0$ and

$\int_1^{\infty} \frac{dx}{x^p} = \frac{1}{p-1}$; if $p < 1$ the integral is divergent because in the latter case we have $\lim_{x \rightarrow \infty} x^{-p+1} = \infty$. For $p = 1$ the integral is also

divergent: $\int_1^{\infty} \frac{dx}{x} = \ln x \Big|_1^{\infty} = \infty$. Therefore the series is convergent

for $p > 1$ and divergent for $p \leq 1$.

By the hypothesis, the expression in each bracket is positive; since the number of these brackets increases together with m the sequence s_{2m} is *increasing*. On the contrary, we similarly show that the sequence of partial sums s_{2m+1} with odd indices is *decreasing* since

$$s_{2m+1} = u_1 - [(u_2 - u_3) + (u_4 - u_5) + \dots + (u_{2m} - u_{2m+1})]$$

and the sum in the square brackets increases together with m , which is clear from the above argument.

Thus,

$$s_2 < s_4 < s_6 < \dots \text{ and } s_1 > s_3 > s_5 > \dots$$

It is readily seen that any sum with even number is less than any sum with odd number:

$$s_{2m} < s_{2k+1}$$

Indeed, if $m < k$ then

$$s_{2m} < s_{2k} = s_{2k+1} - u_{2k+1} < s_{2k+1}$$

and if $m > k$ (or $m = k$) then

$$s_{2m} = s_{2m+1} - u_{2m+1} < s_{2m+1} < s_{2k+1}$$

In Fig. 247 the sequence of partial sums is represented geometrically. The sequence of sums with even indices is increasing and bounded above; hence, it possesses a limit $\lim_{m \rightarrow \infty} s_{2m} = s$.

We also have $s_{2m+1} = s_{2m} + u_{2m+1}$ and $u_{2m+1} \rightarrow 0$ for $m \rightarrow \infty$; therefore,

$$\lim_{m \rightarrow \infty} s_{2m+1} = \lim_{m \rightarrow \infty} s_{2m} = s$$

that is the sums with odd numbers (forming a decreasing sequence) tend to the same limit s .

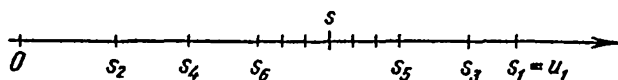


Fig. 247

Fig. 247 also clearly indicates that the partial sums approach their limit s so that those with even numbers are less than s while those with odd numbers are greater than s ; in particular, the sum s of the series is less than u_1 (since we have $s_1 = u_1$).

If the whole series (*) is taken with the minus sign we should take the reflection of Fig. 247 about the origin; in this case $|s| < u_1$.

The remainder

$$r_n = \pm (u_{n+1} - u_{n+2} + \dots)$$

The greater the number p , the more rapidly the p series converges. For example, if $p=3$ we have $r_n < \frac{1}{2n^2}$, and the same accuracy of $\epsilon=0.001$ is achieved for $n=24$. If $p=4$ then $r_n < \frac{1}{3n^3}$, and the same precision is obtained for $n=7$.

Thus, we see that the faster the terms of the series tend to zero, the more rapid the convergence of the series.

In special courses on approximate calculations the reader can find various methods for *accelerating* the convergence of a series. These techniques make it sometimes possible to find an accurate approximation to the sum of a slowly convergent series with a comparatively small amount of calculations. By the way, some of these techniques make use of the fact that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \approx 1.645$ (see Sec. 201).

We also note that for the time being it is not known whether the sum of the series $\sum_{n=1}^{\infty} \frac{1}{n^3}$ is a rational or irrational number; its approximate value is 1.202. For $p=4$ the sum of the p series is known: $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90} \approx 1.082$.

186. Series with Arbitrary Terms. Absolute Convergence.

1. Alternating series. We now proceed to investigate series with terms of arbitrary signs and begin with the so-called (*alternating series*) whose terms are alternately positive and negative. Such a series can be written as

$$\pm (u_1 - u_2 + u_3 - \dots) \quad (*)$$

where $u_1, u_2, u_3, \dots$ are *positive* numbers. We shall prove a simple *sufficient* test for convergence of an alternating series:

Theorem (Leibniz' Test). If the absolute values of the terms of an alternating series form a monotone decreasing sequence tending to zero, that is, if the terms of series (*) satisfy the inequalities $u_1 > u_2 > \dots$ and the general term u_n tends to zero, the series is convergent; the absolute value of its sum is less than u_1 , and the absolute value of the remainder r_n of the series is less than that of the first discarded term: $|r_n| < u_{n+1}$.

Proof. For definiteness, let us take the whole expression (*) with the plus sign:

$$u_1 - u_2 + u_3 - \dots - u_{2m} + u_{2m+1} - \dots$$

Consider the sequence of partial sums s_{2m} of the series with even indices $n=2m$. Such a sum can be written as

$$s_{2m} = (u_1 - u_2) + (u_3 - u_4) + \dots + (u_{2m-1} - u_{2m})$$

By the hypothesis, the expression in each bracket is positive; since the number of these brackets increases together with m the sequence s_{2m} is *increasing*. On the contrary, we similarly show that the sequence of partial sums s_{2m+1} with odd indices is *decreasing* since

$$s_{2m+1} = u_1 - [(u_2 - u_3) + (u_4 - u_5) + \dots + (u_{2m} - u_{2m+1})]$$

and the sum in the square brackets increases together with m , which is clear from the above argument.

Thus,

$$s_2 < s_4 < s_6 < \dots \text{ and } s_1 > s_3 > s_5 > \dots$$

It is readily seen that any sum with even number is less than any sum with odd number:

$$s_{2m} < s_{2k+1}$$

Indeed, if $m < k$ then

$$s_{2m} < s_{2k} = s_{2k+1} - u_{2k+1} < s_{2k+1}$$

and if $m > k$ (or $m = k$) then

$$s_{2m} = s_{2m+1} - u_{2m+1} < s_{2m+1} < s_{2k+1}$$

In Fig. 247 the sequence of partial sums is represented geometrically. The sequence of sums with even indices is increasing and bounded above; hence, it possesses a limit $\lim_{m \rightarrow \infty} s_{2m} = s$.

We also have $s_{2m+1} = s_{2m} + u_{2m+1}$ and $u_{2m+1} \rightarrow 0$ for $m \rightarrow \infty$; therefore,

$$\lim_{m \rightarrow \infty} s_{2m+1} = \lim_{m \rightarrow \infty} s_{2m} = s$$

that is the sums with odd numbers (forming a decreasing sequence) tend to the same limit s .

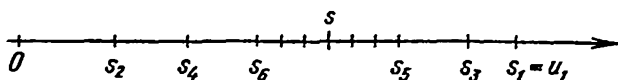


Fig. 247

Fig. 247 also clearly indicates that the partial sums approach their limit s so that those with even numbers are less than s while those with odd numbers are greater than s ; in particular, the sum s of the series is less than u_1 (since we have $s_1 = u_1$).

If the whole series (*) is taken with the minus sign we should take the reflection of Fig. 247 about the origin; in this case $|s| < u_1$.

The remainder

$$r_n = \pm (u_{n+1} - u_{n+2} + \dots)$$

of series (*) is a series satisfying all the conditions of Leibniz' test. Therefore, the sum r_n of this series is less, in its absolute value, than the first term in the parentheses, i.e. $|r_n| < u_{n+1}$, which is what we had to prove. The proof of Leibniz' theorem has been completed.

Leibniz' test, when applicable, not only makes it possible to investigate the convergence of the series but also to estimate the error appearing when the sum of the series is replaced by its partial sum with given number n .

For instance, Leibniz' test immediately shows that the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n+1}}{n} + \dots$$

is convergent; as will be shown in Sec. 194, its sum is equal to $\ln 2$. According to the above, the limiting absolute error of the approximate equality $\ln 2 \approx 1 - \frac{1}{2} + \frac{1}{3} - \dots + \frac{(-1)^n}{n-1}$ is equal to $\frac{1}{n}$, which indicates that the series is slowly convergent.

II. Absolute convergence. Now we proceed to series whose terms can have arbitrary signs. We shall confine ourselves to the following important test:

Sufficient Test for Convergence. If the series whose terms are the absolute values of the terms of a given series is convergent the latter series is also convergent.

Proof. Denote by s_n the sum of the first n terms of the given series

$$u_1 + u_2 + \dots + u_n + \dots$$

and by s_n^+ and s_n^- , respectively, the sums of the absolute values of all positive and all negative summands among these n terms. Then

$$s_n = s_n^+ - s_n^- \quad \text{and} \quad \sigma_n = s_n^+ + s_n^-$$

where

$$\sigma_n = |u_1| + |u_2| + \dots + |u_n|$$

Since, by the hypothesis, σ_n possesses a limit (denote it σ), and s_n^+ and s_n^- are positive increasing functions of n satisfying the conditions $s_n^+ \leq \sigma_n < \sigma$ and $s_n^- \leq \sigma_n < \sigma$, these functions also have limits. It follows that $s_n = s_n^+ - s_n^-$ also tends to a limit as $n \rightarrow \infty$, which is what we set out to prove.

This test does not provide a *necessary* condition for convergence:

a series $\sum_{n=1}^{\infty} u_n$ may be convergent when the series $\sum_{n=1}^{\infty} |u_n|$ is divergent. For instance, as has been shown, the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n+1}}{n} + \dots$$

is convergent while, as we know, the (harmonic) series

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n} + \dots$$

formed of the absolute values of the terms of the former series is divergent.

Definition. A series for which the absolute values of its terms form a convergent series is said to be *absolutely convergent*.

If a given series is convergent whereas the series composed of the absolute values of its terms is divergent the given series is said to be *conditionally convergent*.

An instance of a conditionally convergent series is $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n+1}}{n}$.

The distinction between absolute and conditional (*not absolute*) convergence is extremely important. It turns out that some properties of finite sums can only be extended to absolutely convergent series whereas conditionally convergent series do not possess these properties.

We shall not treat these questions in detail and only confine ourselves to the remark that absolutely convergent series, like ordinary sums of finite number of summands, possess the *commutativity property*: any rearrangement of the terms of an absolutely convergent series does not affect its sum, that is the new series remains convergent and has the same sum as before. Conditionally convergent series do not possess this property: it can be shown that, given any number, it is always possible to rearrange an (infinite) number of terms of a given conditionally convergent series so that its sum becomes equal to the given number (moreover, it is also possible to make the partial sums of the series tend to $+\infty$ or $-\infty$ by rearranging its terms in an appropriate manner).

As an example, let us take the conditionally convergent series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots$$

and show (without rigorous proof) that its sum can be changed by rearranging its terms. To this end, let us rearrange the terms in such a way that every positive term is followed by exactly two negative terms:

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \dots$$

On adding together every positive term and the next negative term we obtain

$$\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \dots$$

The terms of the resultant series are half the corresponding terms of the original series. By Property 1, Sec. 182, it follows that the latter series is also convergent and its sum is half the sum of the former series. We see that the sum of the series became half as great after a mere change of the order of its terms!

Another important property of absolutely convergent series is related to the operation of multiplication.

The (*Cauchy*) *product* of two convergent series

$$s' = u_1 + u_2 + \dots + u_n + \dots \quad (1)$$

and

$$s'' = v_1 + v_2 + \dots + v_n + \dots \quad (2)$$

is the series formed of all possible pairwise products of the terms of the given series arranged as

$$(u_1v_1) + (u_1v_2 + u_2v_1) + (u_1v_3 + u_2v_2 + u_3v_1) + \dots \\ \dots + (u_1v_n + u_2v_{n-1} + \dots + u_{n-1}v_2 + u_nv_1) + \dots \quad (3)$$

To understand the structure of this series note that each group of terms taken into parentheses has a constant sum of the numbers of the factors: for the first group this sum is equal to 2, for the second to 3, for the third to 4, ..., for the n th to $n+1$, etc.

The following theorem which we state without proof describes the multiplication property of absolutely convergent series:

If series (1) and (2) are absolutely convergent their product is also an absolutely convergent series and its sum is equal to the product of the sums of series (1) and (2):

$$s = s' \cdot s''$$

§ 2. Functional Series

187. General Definitions. In § 2 we shall study series whose terms are not numbers but functions:

$$\left(u_1(x) + u_2(x) + \dots + u_n(x) + \dots \right) \quad (*)$$

An expression of type (*) is termed a *functional series*. The functions $u_n(x)$ are supposed to be defined and continuous in one and the same interval (finite or infinite).¹

Series (*) may converge for some values of x and diverge for the other values. A point $x = x_0$ for which the number series

$$u_1(x_0) + u_2(x_0) + \dots + u_n(x_0) + \dots$$

is convergent is called a *point of convergence of series (*)*. The set of all the points of convergence is the *domain of convergence of the series*; we also say that the *series is convergent in this domain*. As a rule, the domain of convergence of a functional series

is an interval of the Ox -axis. For example, the series

$$1 + x + x^2 + \dots + x^n + \dots$$

is convergent in the interval $(-1, 1)$ since for any value $x = x_0$ belonging to this interval the corresponding number series is a convergent geometric series with common ratio x_0 ($|x_0| < 1$). For $|x| \geq 1$ this series is divergent. In the general case the determination of the domain of convergence of a given functional series is an extremely difficult problem.

The sum of a functional series is a function of the variable x defined in the domain of convergence of the series; we denote it $S(x)$:

$$S(x) = u_1(x) + u_2(x) + \dots + u_n(x) + \dots$$

For instance, in the above example the sum of the series is equal to $\frac{1}{1-x}$; this function is of course the sum of the series only within the interval $(-1, 1)$ serving as the domain of convergence of the series although the function $\frac{1}{1-x}$ itself is defined for all $x \neq 1$.

Let us denote by $s_n(x)$ the sum of the first n terms of series (*) (the n th partial sum) and by $r_n(x)$ the remainder after n terms. If the series is convergent for a given value of x then

$$\lim_{n \rightarrow \infty} s_n(x) = S(x) \text{ and } \lim_{n \rightarrow \infty} r_n(x) = 0$$

As is well known, the sum of a finite number of continuous functions is again a continuous function, and the integral of the sum of functions is equal to the sum of the integrals of the constituent functions. Similarly, the derivative of the sum of a finite number of differentiable functions is equal to the sum of the derivatives of the summands. Therefore, in the theory of functional series it is naturally to pose the question as to whether these properties are extended to sums of an infinite number of summands, that is to functional series. The example below demonstrates that in the general case an arbitrary functional series does not necessarily possess these properties.

Let us take the series

$$\frac{x^2}{1+x^2} + \frac{x^2}{(1+x^2)^2} + \dots + \frac{x^2}{(1+x^2)^n} + \dots$$

All the terms and the sum of the series turn into zero for $x=0$. For $x \neq 0$ we obtain a convergent geometric series since the common ratio is $q = \frac{1}{1+x^2} < 1$. For $x \neq 0$ the sum of the series is

equal to

$$\frac{\left(\frac{x^2}{1+x^2}\right)}{1-\frac{1}{1+x^2}}=1$$

We see that the sum $S(x)$ of the series is a discontinuous function with a jump discontinuity at the point $x=0$:

$$S(x)=1 \text{ if } x \neq 0 \text{ and } S(0)=0$$

To find out what conditions guarantee the possibility of extending the properties of finite sums of functions to functional series we need the following definition.

Definition. A functional series

$$u_1(x) + u_2(x) + \dots + u_n(x) + \dots \quad (*)$$

is said to be *regularly convergent* in a domain D belonging to the domain of convergence if for all the points of the domain D the absolute values of the terms of series $(*)$ do not exceed the corresponding terms of a convergent positive numerical series.

(Compare this definition with the definition of a regularly convergent improper integral dependent on a parameter stated in Sec: 137.)

The definition means that for all the points of the domain D the inequalities $|u_n(x)| \leq M_n$, $n=1, 2, \dots$, must hold where M_n is the n th term of a convergent positive numerical series $M_1 + M_2 + \dots + M_n + \dots$; the latter series is spoken of as a *dominant series* of series $(*)$ and series $(*)$ is referred to as the *dominated series*.

For instance, the series

$$\sin x + \frac{\sin 2x}{2^2} + \frac{\sin 3x}{3^2} + \dots + \frac{\sin nx}{n^2} + \dots$$

is regularly convergent in any interval of the x -axis since

$$\left| \frac{\sin nx}{n^2} \right| \leq \frac{1}{n^2} \text{ and the series } \sum_{n=1}^{\infty} \frac{1}{n^2} \text{ is convergent.}$$

Let us briefly discuss a more general notion of a *uniformly convergent series*. Suppose that functional series $(*)$ is convergent at a point x_1 . This means that $\lim_{n \rightarrow \infty} r_n(x_1) = 0$; therefore, given an arbitrary number $\varepsilon > 0$, there is a number N_1 such that the inequality $|r_n(x_1)| < \varepsilon$ is fulfilled for all $n > N_1$. Let us take another point x_2 belonging to the domain of convergence of the series; for this point we also have $\lim_{n \rightarrow \infty} r_n(x_2) = 0$. If we take the same number ε the inequality $|r_n(x_2)| < \varepsilon$ is of course fulfilled

for the values of n exceeding a certain number N_2 , but it may turn out that $N_2 > N_1$. That is why the numbers N_1 and N_2 are supplied with different subscripts.

If the domain of convergence consisted of only two points we could replace N_1 and N_2 by one number N equal to the greatest of the two numbers N_1 and N_2 : if $n > N$ then $n > N_1$ and $n > N_2$, and both inequalities $|r_n(x_1)| < \varepsilon$ and $|r_n(x_2)| < \varepsilon$ hold simultaneously. The same of course applies to the case of the domain of convergence consisting of an arbitrary finite number of points.

But if the domain of convergence is an interval including an *infinity* of points then, although for every point x there is a number $N(\varepsilon, x)$ (we write $N(\varepsilon, x)$ to stress that this number may depend both on ε and on x) such that for $n > N(\varepsilon, x)$ the remainder satisfies the inequality $|r_n(x)| < \varepsilon$, a number $N(N = N(\varepsilon))$ does not depend on x) common for *all the points* of the domain of convergence does not necessarily exist. In this connection the meaning of the following definition becomes clear:

Definition. Functional series (*) is said to be *uniformly convergent* in a domain D belonging to its domain of convergence if, given any number $\varepsilon > 0$, there is a number $N(\varepsilon)$ dependent solely on ε and independent of x such that for $n > N(\varepsilon)$ the inequality $|r_n(x)| < \varepsilon$ is fulfilled for all the points of the domain D .

It can easily be shown that *if a functional series is regularly convergent in a domain D it is also uniformly convergent*. In fact, the definition of regular convergence implies that for any x belonging to the domain D we have

$$|r_n(x)| = |u_{n+1}(x) + \dots| \leq |u_{n+1}(x)| + \dots \leq M_{n+1} + \dots = r_n$$

where r_n is the remainder of the convergent numerical series $M_1 + M_2 + \dots + M_n + \dots$. The value of r_n is independent of x and is less than any number $\varepsilon > 0$ provided that $n > N$ where N is dependent solely on ε . Then, for the same values of n , we have $|r_n(x)| < \varepsilon$, which shows that the functional series is uniformly convergent. But it is possible to construct examples (we do not present them here) showing that there are uniformly convergent functional series possessing no (convergent) dominant series; such uniformly convergent series are thus not regularly convergent.

For the sake of simplicity, we shall speak of regularly convergent series although all the theorems stated in Sec. 188 hold for uniformly convergent series as well.

188. Properties of Regularly Convergent Functional Series. In this section we shall state without proof some fundamental theorems on regularly convergent functional series which elucidate the question posed in Sec. 187 (on the possibility of extending the properties of sums of a finite number of functions to func-

tional series). In all the theorems below we assume that the domain of regular convergence of the series in question is an interval of the axis Ox .

Theorem I. If a series formed of continuous functions is regularly convergent in a domain D its sum is also a continuous function in that domain.

As was shown, the series $\sin x + \frac{\sin 2x}{2^2} + \dots + \frac{\sin nx}{n^2} + \dots$ is regularly convergent in any interval; hence, its sum $S(x)$ is a continuous function throughout Ox .

Practical determination of this function is an essentially more difficult problem; it is even unknown whether this function is elementary or not. But we can always find an approximation to the value of this function for any point of the domain of convergence by taking the sum of a sufficiently large finite number of terms of the series, that is by computing $S_n(x)$ for a sufficiently large n . It should be stressed that, because of the regular convergence, the limiting absolute error appearing in such an approximation is one and the same for all the values of x since

$$|r_n(x)| = \left| \frac{\sin(n+1)x}{(n+1)^2} + \dots \right| \leq \left| \frac{\sin(n+1)x}{(n+1)^2} \right| + \dots \leq \frac{1}{(n+1)^2} + \dots = r_n$$

where r_n is the remainder of the series with the general term $\frac{1}{(n+1)^2}$ dependent solely on n and independent of x . The same considerations of course apply to any regularly convergent series. (The reader is advised to try to extend this remark to uniformly convergent series; see Sec. 187.)

Thus, the theorem says that the sum of a regularly convergent series whose terms are continuous functions is a continuous function in the domain of regular convergence. In theorem II we shall extend the simplest property of the integral to such series: the definite integral of a sum of a finite number of continuous functions is equal to the sum of the integrals of the summands, this property applying to regularly convergent series as well. We shall assume (without further stipulations) that the interval of integration $[a, b]$ always belongs to the domain of regular convergence of the series in question.

Theorem II. If a series composed of continuous functions is regularly convergent the integral of the sum of the series is equal to the sum of the series whose terms are the integrals of these functions:

$$\int_a^b S(x) dx = \int_a^b u_1(x) dx + \int_a^b u_2(x) dx + \dots + \int_a^b u_n(x) dx + \dots \quad (*)$$

In other words, briefly:

A regularly convergent series can be integrated term-by-term.

The fact that the numerical series formed of the integrals of the terms of the given series is convergent is proved quite simply. The definition of regular convergence implies that for the terms of the given series we have $|u_n(x)| \leq M_n$ where $M_n > 0$ is the n th term of a convergent numerical series. By the theorem on the estimation of the definite integral (see Sec. 90), we have

$$\left| \int_a^b u_n(x) dx \right| \leq \int_a^b |u_n(x)| dx \leq M_n(b-a)$$

Since, according to Property 1 stated in Sec. 182, the series with general term $M_n(b-a)$ is convergent, the application of the results of Sec. 186, II indicates that the series formed of the integrals is convergent (even absolutely convergent). The proof of the second part of the theorem, that is of the fact that the left-hand and the right-hand sides of formula (*) are equal is omitted here.

Theorem II also applies to integrals with variable upper limit of integration:

$$\int_a^x S(x) dx = \int_a^x u_1(x) dx + \int_a^x u_2(x) dx + \dots + \int_a^x u_n(x) dx + \dots$$

where $a \leq x \leq b$. The expression on the right is a functional series; using again the theorem on the estimation of the integral

(i.e. the inequality $\left| \int_a^x u_n(x) dx \right| \leq M_n(x-a) \leq M_n(b-a)$) we readily see that this series, like the given series, is regularly convergent.

The theorem on termwise differentiation of a functional series is more complicated and involves some additional requirements.

Theorem III. If a series $u_1(x) + u_2(x) + \dots + u_n(x) + \dots$ formed of functions possessing continuous derivatives converges in a domain D , its sum being $S(x)$, and if the series composed of the derivatives $u'_n(x)$ is regularly convergent in this domain, the derivative $S'(x)$ of the sum is equal to the sum of the series of the derivatives:

$$S'(x) = u'_1(x) + u'_2(x) + \dots + u'_n(x) + \dots$$

In other words, briefly:

If the series composed of the derivatives of the terms of a convergent functional series is regularly convergent the given functional series can be differentiated term-by-term.

Note that the regular convergence of the given series and also the differentiability of its sum are not assumed in the statement of this theorem but are established in the course of the proof. But the condition that the series composed of the derivatives is regularly convergent is essential and must necessarily be checked; if this condition is violated the theorem may fail. For instance, the series $\sin x + \frac{\sin 2^2 x}{2^2} + \dots + \frac{\sin n^2 x}{n^2} + \dots$ is regularly convergent

(as its dominant series we can take the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$) and its sum is a continuous function. At the same time the series $\cos x + 2 \cos 2^2 x + \dots + n \cos n^2 x + \dots$ formed of the derivatives is divergent since, as $n \rightarrow \infty$, its general term does not tend to zero for any value of x . Consequently, the theorem proves inapplicable.

In this course we shall limit ourselves to studying only two important classes of functional series: the so-called *power series* and *trigonometric series*.

§ 3. Power Series

189. Abel's Theorem. Interval of Convergence and Radius of Convergence. We begin with the definition of a power series.

Definition. A *power series* is a functional series of the form

$$a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + \dots$$

whose terms are the products of constant factors $a_0, a_1, a_2, \dots, a_n, \dots$ by integral powers of the difference $x - x_0$.

The constants $a_0, a_1, a_2, \dots, a_n, \dots$ are called the *coefficients of the power series*. Unless otherwise stated, the coefficients will be assumed to be real.

In particular, if $x_0 = 0$ the power series is arranged in ascending powers of x :

$$a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$$

We shall confine ourselves to studying series of the latter type because any power series can be brought to this type with the aid of the substitution $x - x_0 = x'$.

For the sake of convenience, we shall call $a_n x^n$ the n th term of the power series $a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$ (although it occupies the $(n+1)$ th place). The constant term a_0 of the series will be referred to as its zeroth term.

Power series possess a number of specific properties which we shall study here.

For series

$$a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots \quad (*)$$

we shall prove the following fundamental theorem on which the whole theory of power series is based:

Abel's Theorem. If power series (*) converges at a point $x_0 \neq 0$ it is absolutely convergent in the interval $(-|x_0|, |x_0|)$, i.e. for every value of x satisfying the condition $|x| < |x_0|$.

Proof. The series $\sum_{n=0}^{\infty} a_n x_0^n$ being convergent, its general term tends to zero: $a_n x_0^n \rightarrow 0$; therefore the set of values of these terms is bounded, that is there exists a positive number M such that $|a_n x_0^n| < M$ for any n . Let us write down series (*) as

$$a_0 + a_1 x_0 \left(\frac{x}{x_0}\right) + a_2 x_0^2 \left(\frac{x}{x_0}\right)^2 + \dots + a_n x_0^n \left(\frac{x}{x_0}\right)^n + \dots$$

and form the series of the absolute values of its terms:

$$|a_0| + |a_1 x_0| \left|\frac{x}{x_0}\right| + |a_2 x_0^2| \left|\frac{x}{x_0}\right|^2 + \dots + |a_n x_0^n| \left|\frac{x}{x_0}\right|^n + \dots$$

By the inequality proved, every term of the latter series is less than the corresponding term of the geometric series with ratio $\left|\frac{x}{x_0}\right|$:

$$M + M \left|\frac{x}{x_0}\right| + M \left|\frac{x}{x_0}\right|^2 + \dots + M \left|\frac{x}{x_0}\right|^n + \dots$$

For $|x| < |x_0|$ we have $\left|\frac{x}{x_0}\right| < 1$, and the geometric series is convergent; therefore the series of the absolute values of the terms of the given power series is also convergent, and hence series (*) is itself absolutely convergent. The theorem has been proved.

Note that the inequality $|a_n x^n| < |a_n x_0^n|$ does not enable us to use directly the comparison test since the conditions of the theorem do not indicate that the given series is absolutely convergent at the point x_0 itself.

Corollary. If power series (*) diverges for $x = x_0$ it also diverges for any x exceeding x_0 in its absolute value, that is for $|x| > |x_0|$.

If the series were convergent for such a value of x , Abel's theorem would imply its convergence for the values of the argument smaller than $|x|$ in their absolute values and, in particular, for $x = x_0$, which contradicts the hypothesis.

Let us proceed to investigate the domain of convergence of series (*). There are three possible cases here:

(1) The domain of convergence consists of one point $x = 0$; in other words, the series is divergent for all the points except $x = 0$.

Such a case can be demonstrated by the example of the series

$$1 + x + 2^2 x^2 + \dots + n^n x^n + \dots$$

Indeed, for any fixed $x \neq 0$ we have $|nx| > 1$ beginning with a

sufficiently large value of n whence follows the inequality $|n^n x^n| > 1$ indicating that the general term of the series does not tend to zero.

(2) *The domain of convergence consists of all the points of the x -axis, that is the series is convergent for all x 's (such a series is said to be permanently convergent).*

An example of such a series is

$$1 + x + \frac{x^2}{2^2} + \dots + \frac{x^n}{n^n} + \dots$$

Indeed, for any x we have $\left|\frac{x}{n}\right| < 1$ beginning with a sufficiently large n . Since

$$\left|\frac{x}{n+1}\right|^{n+1} < \left|\frac{x}{n}\right|^{n+1}, \quad \left|\frac{x}{n+2}\right|^{n+2} < \left|\frac{x}{n}\right|^{n+2}, \text{ etc.}$$

the absolute values of the terms of the series become, beginning with the number n , less than the terms of the convergent geometric series with common ratio $\left|\frac{x}{n}\right| < 1$. Consequently, the given series is convergent for any x .

(3) *The domain of convergence consists of more than one point, and there are points of the axis Ox not belonging to the domain of convergence.*

For instance, the series

$$1 + x + x^2 + \dots + x^n + \dots$$

(the geometric series with common ratio x) is convergent for $|x| < 1$ and divergent for $|x| \geq 1$.

In the latter case the Ox axis contains both points of convergence and points of divergence of the series.

Abel's theorem and its corollary imply that *all* the points of convergence are *not farther* from the origin than any point of divergence. It is also clear that the points of convergence entirely cover a whole interval with centre at the origin.

Thus, we can say that *for every power series possessing both points of convergence and points of divergence there is a positive number R such that for all x 's whose absolute values are less than R ($|x| < R$) the series converges absolutely and for all x 's exceeding R in their absolute values ($|x| > R$) it diverges.*

As for the values $x = R$ and $x = -R$ of the argument, there can be various possibilities in these cases: the series may converge at both points or at one of them or at neither. Moreover, the convergence at these points may be absolute or conditional. Later on we shall consider examples illustrating these possibilities.

Definition. The number R such that power series (*) is convergent for all x satisfying the condition $|x| < R$ and divergent for

all x satisfying the inequality $|x| > R$ is called the *radius of convergence* of the power series. The interval $(-R, R)$ is referred to as the *interval of convergence*.

Let us agree that $R=0$ for the series divergent for all the values of x except $x=0$ and that $R=+\infty$ (or, simply, $R=\infty$) for the permanently convergent series (i.e. for the series convergent for all the x 's).

All that has been said also applies to power series of the general form

$$a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \dots + a_n(x-x_0)^n + \dots$$

with the only distinction that in the general case the centre of the interval of convergence is not at the point $x=0$ but at the point $x=x_0$, the interval of convergence being (x_0-R, x_0+R) .

We shall present a method for determining the radius of convergence of a power series. To this end, note that it is sufficient to find the radius of convergence for the series

$$|a_0| + |a_1||x| + \dots + |a_n||x|^n + \dots \quad (**)$$

composed of the absolute values of the terms of series (*) since the intervals of convergence of series (*) and (**) coincide.

Let us apply D'Alembert's test to the positive series (**). Suppose that the limit of the ratio of two successive terms of this series exists and that this limit has been found. In contrast to the case of a numerical series, this limit may depend on $|x|$. Then, for the values of $|x|$ for which this limit is *less than* 1 the series is convergent and for the values of $|x|$ for which the limit *exceeds* 1 it is divergent. It follows that the value of $|x|$ for which this limit is equal to unity specifies the radius of convergence of power series (*).

It may turn out that this limit is equal to zero for *all* $x \neq 0$. This means that series (*) converges for all the values of x and its radius of convergence is equal to infinity ($R=\infty$).

On the contrary, if the limit is equal to infinity for all $x \neq 0$ the series is everywhere divergent (except at the point $x=0$) and its radius of convergence is equal to zero.

Let us demonstrate what has been said by concrete examples.

Examples. (1) Let us determine the radius of convergence of the series

$$1 + x + \frac{1}{2!}x^2 + \dots + \frac{1}{n!}x^n + \dots$$

We have

$$\left| \frac{u_{n+1}}{u_n} \right| = \frac{|x|^{n+1} n!}{(n+1)! |x|^n} = \frac{|x|}{n+1} \rightarrow 0$$

for any value of x . Hence, the series is convergent for all x 's, that is $R=\infty$.

sufficiently large value of n whence follows the inequality $|n^n x^n| > 1$ indicating that the general term of the series does not tend to zero.

(2) *The domain of convergence consists of all the points of the x -axis, that is the series is convergent for all x 's (such a series is said to be permanently convergent).*

An example of such a series is

$$1 + x + \frac{x^2}{2^2} + \dots + \frac{x^n}{n^n} + \dots$$

Indeed, for any x we have $\left|\frac{x}{n}\right| < 1$ beginning with a sufficiently large n . Since

$$\left|\frac{x}{n+1}\right|^{n+1} < \left|\frac{x}{n}\right|^{n+1}, \quad \left|\frac{x}{n+2}\right|^{n+2} < \left|\frac{x}{n}\right|^{n+2}, \quad \text{etc.}$$

the absolute values of the terms of the series become, beginning with the number n , less than the terms of the convergent geometric series with common ratio $\left|\frac{x}{n}\right| < 1$. Consequently, the given series is convergent for any x .

(3) *The domain of convergence consists of more than one point, and there are points of the axis Ox not belonging to the domain of convergence.*

For instance, the series

$$1 + x + x^2 + \dots + x^n + \dots$$

(the geometric series with common ratio x) is convergent for $|x| < 1$ and divergent for $|x| \geq 1$.

In the latter case the Ox axis contains both points of convergence and points of divergence of the series.

Abel's theorem and its corollary imply that *all* the points of convergence are *not farther* from the origin than any point of divergence. It is also clear that the points of convergence entirely cover a whole interval with centre at the origin.

Thus, we can say that *for every power series possessing both points of convergence and points of divergence there is a positive number R such that for all x 's whose absolute values are less than R ($|x| < R$) the series converges absolutely and for all x 's exceeding R in their absolute values ($|x| > R$) it diverges.*

As for the values $x = R$ and $x = -R$ of the argument, there can be various possibilities in these cases: the series may converge at both points or at one of them or at neither. Moreover, the convergence at these points may be absolute or conditional. Later on we shall consider examples illustrating these possibilities.

Definition. The number R such that power series (*) is convergent for all x satisfying the condition $|x| < R$ and divergent for

all x satisfying the inequality $|x| > R$ is called the *radius of convergence* of the power series. The interval $(-R, R)$ is referred to as the *interval of convergence*.

Let us agree that $R=0$ for the series divergent for all the values of x except $x=0$ and that $R=+\infty$ (or, simply, $R=\infty$) for the permanently convergent series (i.e. for the series convergent for all the x 's).

All that has been said also applies to power series of the general form

$$a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \dots + a_n(x-x_0)^n + \dots$$

with the only distinction that in the general case the centre of the interval of convergence is not at the point $x=0$ but at the point $x=x_0$, the interval of convergence being (x_0-R, x_0+R) .

We shall present a method for determining the radius of convergence of a power series. To this end, note that it is sufficient to find the radius of convergence for the series

$$|a_0| + |a_1||x| + \dots + |a_n||x|^n + \dots \quad (**)$$

composed of the absolute values of the terms of series (*) since the intervals of convergence of series (*) and (**) coincide.

Let us apply D'Alembert's test to the positive series (**). Suppose that the limit of the ratio of two successive terms of this series exists and that this limit has been found. In contrast to the case of a numerical series, this limit may depend on $|x|$. Then, for the values of $|x|$ for which this limit is *less than* 1 the series is convergent and for the values of $|x|$ for which the limit *exceeds* 1 it is divergent. It follows that the value of $|x|$ for which this limit is equal to unity specifies the radius of convergence of power series (*).

It may turn out that this limit is equal to zero for *all* $x \neq 0$. This means that series (*) converges for all the values of x and its radius of convergence is equal to infinity ($R=\infty$).

On the contrary, if the limit is equal to infinity for all $x \neq 0$ the series is everywhere divergent (except at the point $x=0$) and its radius of convergence is equal to zero.

Let us demonstrate what has been said by concrete examples.

Examples. (1) Let us determine the radius of convergence of the series

$$1 + x + \frac{1}{2!}x^2 + \dots + \frac{1}{n!}x^n + \dots$$

We have

$$\left| \frac{u_{n+1}}{u_n} \right| = \frac{|x|^{n+1}n!}{(n+1)!|x|^n} = \frac{|x|}{n+1} \rightarrow 0$$

for any value of x . Hence, the series is convergent for **all** that is $R=\infty$.

According to the necessary condition for convergence, the general term of this series tends to zero, that is

$$\lim_{n \rightarrow \infty} \frac{x^n}{n!} = 0$$

for any fixed value of x . This fact will be used in what follows.

(2) Let us take the series

$$1 + x + \frac{x^2}{2} + \dots + \frac{x^n}{n} + \dots$$

In this case we have

$$\left| \frac{u_{n+1}}{u_n} \right| = \left| \frac{x^{n+1} n}{(n+1)x^n} \right| = |x| \left| \frac{n}{n+1} \right| \rightarrow |x|$$

Consequently, by D'Alembert's test, the series converges if $|x| < 1$ and diverges if $|x| > 1$. Hence, $R = 1$.

In this example it is easy to investigate the behaviour of the series at the end points of its interval of convergence. For $x = 1$ we obtain the divergent series $1 + 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots$ and for $x = -1$ the conditionally convergent series $1 - 1 + \frac{1}{2} - \frac{1}{3} + \dots$

It should be noted that D'Alembert's test is inapplicable to the *end points* of the interval of convergence (the corresponding limit is equal to unity for these points), and therefore some other means should be used for testing the series for convergence.

(3) Let us investigate the series

$$x - \frac{x^3}{3^2} + \frac{x^5}{5^2} - \dots + (-1)^{n+1} \frac{x^{2n-1}}{(2n-1)^2} + \dots$$

Here we obtain

$$\left| \frac{u_{n+1}}{u_n} \right| = \left| \frac{x^{2n+1} (2n-1)^2}{(2n+1)^2 x^{2n-1}} \right| = |x|^2 \left(\frac{2n-1}{2n+1} \right)^2 \rightarrow |x|^2$$

Hence, the radius of convergence is equal to 1; for $|x| < 1$ the series is convergent and for $|x| > 1$ divergent. Let the reader show that it is absolutely convergent at both end points of the interval of convergence.

(4) Let us determine the interval of convergence of the series

$$\frac{x-1}{1 \cdot 2} + \frac{(x-1)^2}{2 \cdot 2^2} + \frac{(x-1)^3}{3 \cdot 2^3} + \dots + \frac{(x-1)^n}{n \cdot 2^n} + \dots$$

To this end we form the ratio

$$\left| \frac{u_{n+1}}{u_n} \right| = \frac{|x-1|^{n+1} n \cdot 2^n}{(n+1) \cdot 2^{n+1} |x-1|^n} = |x-1| \frac{n}{2(n+1)}$$

and find its limit as $n \rightarrow \infty$:

$$\lim_{n \rightarrow \infty} \left| \frac{u_{n+1}}{u_n} \right| = \frac{|x-1|}{2} \lim_{n \rightarrow \infty} \frac{n}{n+1} = \frac{|x-1|}{2}$$

According to the results obtained, the series is convergent if $\frac{|x-1|}{2} < 1$, that is if $-2 < x-1 < 2$. Therefore the interval of convergence of this series is the interval $(-1, 3)$ with centre at the point $x=1$. Let the reader check that the behaviour of the series at the end points of the interval of convergence is the same as in Example 2.

190. **Properties of Power Series.** Let us consider a power series

$$a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots \quad (*)$$

with radius of convergence R where R can assume any nonnegative value (including $+\infty$). The sum of this series (denote it $S(x)$) is a function defined within the interval of convergence and also at those end points of the interval where the series converges.

To apply the results of Sec. 188 to power series we begin with proving the following lemmas¹.

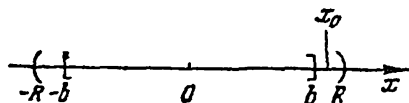


Fig. 248

Lemma I. Power series $(*)$ is regularly convergent in any closed interval $[-b, b]$ lying entirely inside its interval of convergence $(-R, R)$.

Proof. Let us choose a point x_0 such that $b < x_0 < R$ (Fig. 248); this point belongs to the interval of convergence, and Abel's theorem implies that the numerical series

$$|a_0| + |a_1x_0| + \dots + |a_nx_0^n| + \dots$$

is convergent. For any point $x \in [-b, b]$ we have $|x| < x_0$ and, consequently, $|a_nx^n| < |a_nx_0^n|$. It follows from the latter inequality that series $(*)$ is regularly convergent in the interval $[-b, b]$, which we wanted to prove.

Lemma II. The power series formed of the derivatives of the terms of series $(*)$ has the same radius of convergence as the given series.

¹ If Sec. 188 was skipped the reader should proceed directly to the properties of power series stated below.

Proof. The series composed of the derivatives has the form

$$a_1 + 2a_2x + \dots + na_nx^{n-1} + (n+1)a_{n+1}x^n + \dots \quad (**)$$

For the sake of simplicity, let us suppose that the limit of the ratio $\left| \frac{a_{n+1}}{a_n} \right|$ exists¹ as $n \rightarrow \infty$; then the radii of convergence of series (*) and (**) can be found with the aid of D'Alembert's test:

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}x^{n+1}}{a_nx^n} \right| = |x| \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

and

$$\lim_{n \rightarrow \infty} \left| \frac{(n+1)a_{n+1}x^{n+1}}{na_nx^n} \right| = |x| \lim_{n \rightarrow \infty} \frac{n+1}{n} \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = |x| \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$$

which indicates the equality of the limits of the ratios of two successive terms for both series. Hence, their radii of convergence coincide. We also note that series (**) may be divergent at an end point of the interval of convergence even when series (*) is convergent at that end point; the reader can check such a possibility by examining the series $1 + x + \frac{x^2}{2^2} + \dots + \frac{x^n}{n^2} + \dots$.

If we now form the power series of the derivatives of the terms of series (**) the resultant series has the same radius of convergence as the original series and so on. Consequently, *all the power series obtained from series (*) by successive differentiations have the same radius of convergence and, according to Lemma 1, are regularly convergent in any interval lying entirely within the interval of convergence.*

These lemmas and the general properties of regularly convergent series imply the following properties of power series:

1. The sum of a power series is a continuous function in the interval of convergence of the series, i.e. the function

$$S(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n + \dots \quad (-R < x < R)$$

is continuous. Moreover, it can be proved that if a power series is convergent at an end point of its interval of convergence its sum $S(x)$ is continuous as x approaches that end point remaining inside the interval of convergence.

It should also be noted that the analytical expression of the sum of a power series may specify a function whose domain of definition is wider than the interval of convergence of the series. This fact was already mentioned in Sec. 187 in connection with

¹ The lemma also remains true when this limit does not exist but the proof becomes more complicated, and we do not present it here.

the function $\frac{1}{1-x}$ which is the sum of the geometric series

$$1 + x + x^2 + \dots + x^n + \dots$$

convergent to this function for $|x| < 1$. The function $\frac{1}{1-x}$ is defined and continuous everywhere except at the point $x=1$ but it only serves as the sum of the series for $|x| < 1$; for $|x| \geq 1$ the series is divergent, and it is senseless to speak of its sum.

2. A power series can be integrated termwise within its interval of convergence:

$$\begin{aligned} \int_0^x S(x) dx &= \int_0^x a_0 dx + \int_0^x a_1 x dx + \int_0^x a_2 x^2 dx + \dots + \\ &+ \int_0^x a_n x^n dx + \dots = a_0 x + \frac{a_1}{2} x^2 + \frac{a_2}{3} x^3 + \dots + \frac{a_n}{n+1} x^{n+1} + \dots \\ & \qquad \qquad \qquad (-R < x < R) \end{aligned}$$

The new (integrated) series possesses the same radius of convergence as the original series.

3. A power series can be differentiated termwise in its interval of convergence:

$$S'(x) = a_1 + 2a_2 x + \dots + na_n x^{n-1} + \dots \quad (-R < x < R)$$

On differentiating repeatedly we obtain, in succession,

$$\begin{aligned} S''(x) &= 2a_2 + 3 \cdot 2a_3 x + \dots + n(n-1)a_n x^{n-2} + \dots \\ S'''(x) &= 3 \cdot 2a_3 + \dots + n(n-1)(n-2)a_n x^{n-3} + \dots \end{aligned}$$

and so on. Thus,

every power series is infinitely termwise differentiable inside its interval of convergence. The radii of convergence of the differentiated series remain the same as that of the original series.

§ 4. Expanding Functions into Power Series

191. **Taylor's¹ Series.** As was shown, the sum of a power series is a continuous and infinitely differentiable function in the interval of convergence (see Sec. 190). In this section we shall discuss the reverse problem:

Given a function $f(x)$, what conditions guarantee that it can be represented as the sum of a power series?

The properties of power series indicate that if a function $f(x)$ is the sum of a power series it must be infinitely differentiable. But, as will be shown, this condition is not sufficient.

¹ Taylor, B. (1685-1731), an English mathematician.

Our aim is to find out what functions and in what intervals can be represented as sums of power series.

If a function $f(x)$ can be represented as the sum of a power series we say that it is *expanded into the power series*.

The existence of such an expansion is extremely important since it makes it possible to replace, approximately, the given function by the sum of the first several terms of the power series, i.e. by a polynomial. The computation of the values of this function then reduces to the computation of the values of the polynomial, which can be achieved with the aid of the simplest arithmetical operations. The fact that the accuracy of such an approximation will be estimated is also extremely important for practice.

We remind the reader that in Sec. 53 we already dealt with a particular case of polynomial approximation of functions when the differential was used for approximate calculations. Namely, we replaced a function in the vicinity of a given point by a polynomial of the first degree, that is by a linear function. But the theoretical and practical significance of this result was limited since we did not estimate the accuracy of the approximation and the interval within which the approximation was applicable.

The replacement of a function by a simple expression of the type of a polynomial proves extremely convenient in various problems of mathematical analysis such as evaluating integrals, solving differential equations and the like.

Thus, we shall assume that the given function $f(x)$ is infinitely differentiable in a neighbourhood of a point x_0 . Suppose that it is representable as the sum of a power series convergent within an interval containing the point x_0 :

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n + \dots \quad (*)$$

where $a_0, a_1, a_2, \dots, a_n, \dots$ are constant coefficients.

We shall show that the general properties of power series make it possible to find $a_0, a_1, a_2, \dots$, considered as *undetermined coefficients*, by using the known values of the function $f(x)$ and of its derivatives at the point x_0 .

Let us put $x = x_0$ in equality (*). This yields

$$f(x_0) = a_0$$

Now, let us differentiate the power series and substitute $x = x_0$ into the differentiated series

$$f'(x) = a_1 + 2a_2(x - x_0) + \dots + na_n(x - x_0)^{n-1} + \dots$$

This results in

$$f'(x_0) = a_1$$

The repeated differentiation gives us

$$f''(x) = 2a_2 + \dots + n(n-1)a_n(x-x_0)^{n-2} + \dots$$

whence, on substituting $x=x_0$, we obtain

$$f''(x_0) = 2a_2, \text{ i.e. } a_2 = \frac{f''(x_0)}{2!}$$

(to unify the notation in the expressions obtained we write $2!$ instead of 2). Proceeding in this way we obtain, after the n th differentiation,

$$f^{(n)}(x) = n! a_n + \dots$$

where the terms indicated by dots involve the factor $(x-x_0)$ and hence turn into zero for $x=x_0$. Hence, putting $x=x_0$, we receive

$$f^{(n)}(x_0) = n! a_n, \text{ i.e. } a_n = \frac{f^{(n)}(x_0)}{n!}$$

We thus consecutively determine the coefficients $a_0, a_1, a_2, \dots$ of expansion (*):

$$a_0 = f(x_0), \quad a_1 = f'(x_0), \quad a_2 = \frac{f''(x_0)}{2!}, \quad \dots, \quad a_n = \frac{f^{(n)}(x_0)}{n!}, \quad \dots$$

On substituting the expressions found into equality (*) we arrive at the series

$$\begin{aligned} f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \dots \\ \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \dots \end{aligned} \quad (**)$$

called the *Taylor series* of the function $f(x)$.

Definition. Given a function $f(x)$, its *Taylor series* in a neighbourhood of a point x_0 is power series (**) whose coefficients $a_0, a_1, a_2, \dots, a_n, \dots$ are expressed in terms of the function $f(x)$ and its derivatives by the formulas

$$a_0 = f(x_0), \quad a_1 = f'(x_0), \quad a_2 = \frac{f''(x_0)}{2!}, \quad \dots, \quad a_n = \frac{f^{(n)}(x_0)}{n!}, \quad \dots$$

These coefficients are termed the *Taylor coefficients of the function $f(x)$ at x_0* , partial sums of this series being called *Taylor's polynomials*.

Thus, the result we have established is that

if a function $f(x)$ can be expanded into a series in powers of the difference $(x-x_0)$, this series is uniquely determined and necessarily coincides with the Taylor series of this function.

It should be stressed that all our conclusions have been drawn under the assumption that the function $f(x)$ admits of the expansion into a power series.

If this assumption is not made and it is only supposed that the function $f(x)$ is infinitely differentiable, its Taylor series exists

and can be constructed explicitly but the question as to whether it is convergent for any value of x different from x_0 remains open.

Moreover, strange as it may seem, there are cases when the Taylor series constructed for a function $f(x)$ is convergent but not to that function, i.e. its sum is different from $f(x)$.

An example of this kind is the function

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{for } x \neq 0 \\ 0 & \text{for } x = 0 \end{cases}$$

This function is infinitely differentiable throughout Ox and all its derivatives are equal to zero at the point $x=0$. Indeed, we have $f'(x) = \frac{2}{x^3} e^{-\frac{1}{x^2}}$ for $x \neq 0$ and $f'(0) = 0$ since

$$\lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} \frac{e^{-\frac{1}{h^2}}}{h} = \lim_{z \rightarrow \infty} \frac{z}{e^{z^2}} = 0$$

(we have put $\frac{1}{h} = z$). Furthermore, $f''(0) = 0$ since

$$\lim_{h \rightarrow 0} \frac{f'(h) - f'(0)}{h} = \lim_{h \rightarrow 0} \frac{\frac{2}{h^3} e^{-\frac{1}{h^2}}}{h} = \lim_{z \rightarrow \infty} \frac{2z^4}{e^{z^2}} = 0$$

etc. Consequently, all the Taylor coefficients of the function $f(x)$ at $x=0$ are equal to zero, and the corresponding Taylor series consists of terms identically equal to zero and thus converges to the function identically equal to zero but not to the function $f(x)$.

192. Key Condition for Expanding a Function into Taylor's Series. In this section we shall discuss the conditions under which the Taylor series

$$f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \dots$$

formed for a given infinitely differentiable function $f(x)$ converges within an interval, and its sum is exactly equal to $f(x)$ in that interval.

Let $T_n(x)$ denote Taylor's polynomial of the n th degree equal to the n th partial sum of Taylor's series:

$$T_n(x) = f(x_0) + f'(x_0)(x-x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

Putting $R_n(x) = f(x) - T_n(x)$ we can write *Taylor's formula*

$$f(x) = T_n(x) + R_n(x)$$

for the function $f(x)$, $R_n(x)$ being the *remainder after n terms in Taylor's formula*.

The convergence of the Taylor series to the function $f(x)$ at a point x is equivalent to the condition

$$\lim_{n \rightarrow \infty} T_n(x) = f(x)$$

or, which is the same, to the relation

$$\lim_{n \rightarrow \infty} [f(x) - T_n(x)] = \lim_{n \rightarrow \infty} R_n(x) = 0$$

The value of $R_n(x)$ is exactly equal to the error appearing when the function $f(x)$ is replaced by Taylor's polynomial $T_n(x)$.

Before proceeding to the estimation of the remainder we shall discuss a special case. Let us suppose that the function $f(x)$ itself is a polynomial of the n th degree. Then the successive differentiations of the function $f(x)$ each time reduce the degree of the polynomial by unity, and hence, after the n th differentiation we obtain a constant, all the subsequent derivatives being identically equal to zero. Hence, the Taylor series of the polynomial $f(x)$ only contains $n+1$ nonzero terms and thus is again a polynomial of degree n coinciding with $f(x)$ identically. This result can of course be obtained purely algebraically if we simply arrange the given polynomial in ascending power of the difference $x-x_0$.

Thus, we obtain the identity

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n$$

which is Taylor's formula for the polynomial $f(x)$.

Example. Let us arrange the polynomial $f(x) = -3 + x - x^2 + 2x^3$ in ascending powers of the difference $x-1$. Here we have $x_0 = 1$ and

$$\begin{aligned} f(1) &= -1, \quad f'(1) = (1 - 2x + 6x^2)_{x=1} = 5, \\ f''(1) &= (-2 + 12x)_{x=1} = 10 \quad \text{and} \quad f'''(1) = 12 \end{aligned}$$

Hence,

$$-3 + x - x^2 + 2x^3 = -1 + 5(x-1) + 5(x-1)^2 + 2(x-1)^3$$

On opening the parentheses on the right-hand side and combining similar terms we readily check the correctness of the identity obtained.

193. Taylor's Formula. Lagrange's Form of the Remainder in Taylor's Formula. Let $f(x)$ be an infinitely differentiable function for which we want to find out whether it can be expanded into Taylor's series in a neighbourhood of a point x_0 . Representing the function in the form (see Sec. 192)

$$\begin{aligned} f(x) &= T_n(x) + R_n(x) = \\ &= f(x_0) + f'(x_0)(x-x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x) \quad (*) \end{aligned}$$

we obtain Taylor's formula for $f(x)$ with the remainder $R_n(x)$.

The theorem below elucidates the structure of the remainder $R_n(x)$ and makes it possible to answer the question on whether or not $R_n(x)$ tends to zero as n increases indefinitely, that is whether or not the function $f(x)$ can be expanded into Taylor's series.

Lagrange's theorem (On the Remainder after n Terms in Taylor's Formula). If the function $f(x)$ possesses the $(n+1)$ th derivative $f^{(n+1)}(x)$ at all the points of an interval containing the point x_0 the remainder $R_n(x)$ is representable in the form

$$R_n(x) = f^{(n+1)}(\xi) \frac{(x-x_0)^{n+1}}{(n+1)!}$$

for every point x of this interval where ξ is a number lying between x_0 and x .

The latter formula expresses the remainder after n terms in Taylor's formula in Lagrange's form.

Proof. Let us write down the remainder $R_n(x)$ in the form

$$R_n(x) = D_{n+1} \frac{(x-x_0)^{n+1}}{(n+1)!}$$

analogous to that of the general term of series (*) where D_{n+1} is an undetermined magnitude and find the expression of D_{n+1} satisfying the equality

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \dots \\ &\dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + D_{n+1} \frac{(x-x_0)^{n+1}}{(n+1)!} \end{aligned}$$

for every x from the interval indicated in the statement of the theorem.

Let us fix a value x_1 of the variable x ; then D_{n+1} can be written as

$$D_{n+1} = \frac{f(x_1) - f(x_0) - f'(x_0)(x_1-x_0) - \dots - \frac{f^{(n)}(x_0)}{n!}(x_1-x_0)^n}{\frac{(x_1-x_0)^{n+1}}{(n+1)!}}$$

Our aim is to show that D_{n+1} can be represented in the form indicated in the theorem, that is can be written as $f^{(n+1)}(\xi)$ where $\xi \in (x_0, x_1)$. To this end, note that for $n=0$ this follows from Lagrange's formula of finite increments (Sec. 57):

$$D_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} = f'(\xi) \quad \text{where } \xi \in (x_0, x_1)$$

To prove the theorem for any n we construct, like in the proof of Lagrange's theorem on finite increments, an auxiliary function $F(x)$ satisfying the conditions of Rolle's theorem. In this case the

function $F(x)$ has a more complicated expression:

$$F(x) = f(x_1) - f(x) - f'(x)(x_1 - x) - \frac{f''(x)}{2!}(x_1 - x)^2 - \dots \\ \dots - \frac{f^{(n)}(x)}{n!}(x_1 - x)^n - D_{n+1} \frac{(x_1 - x)^{n+1}}{(n+1)!}$$

This expression shows directly that $F(x_1) = 0$. Putting $x = x_0$ and replacing D_{n+1} by its value we see that $F(x_0) = 0$.

Now we find the derivative $F'(x)$:

$$F'(x) = -f'(x) - f''(x)(x_1 - x) + f'(x) - \\ - \frac{f'''(x)}{2!}(x_1 - x)^2 + f''(x)(x_1 - x) - \dots \\ \dots - \frac{f^{(n+1)}(x)}{n!}(x_1 - x)^n + \frac{f^{(n)}(x)}{(n-1)!}(x_1 - x)^{n-1} + D_{n+1} \frac{(x_1 - x)^n}{n!}$$

The derivative $F'(x)$ exists at all the points of the interval since, by the conditions of the theorem, the derivative $f^{(n+1)}(x)$ exists and hence all the derivatives of lower order also exist. In this expression of $F'(x)$ all the terms mutually cancel out except the two printed in bold face¹; on taking the common factor outside the brackets we obtain

$$F'(x) = \frac{(x_1 - x)^n}{n!} [-f^{(n+1)}(x) + D_{n+1}]$$

We have thus verified that the function $F(x)$ satisfies all the conditions of Rolle's theorem and have found its derivative. By Rolle's theorem, there exists a value $x = \xi \in (x_0, x_1)$ such that $F'(\xi) = 0$. Taking this value $x = \xi$ (if there are several such values ξ we take one of them) and substituting it into $F'(x)$ we obtain

$$F'(\xi) = \frac{(x_1 - \xi)^n}{n!} [-f^{(n+1)}(\xi) + D_{n+1}] = 0$$

The first factor on the right is different from zero since $\xi \neq x_1$; hence, there must be $D_{n+1} = f^{(n+1)}(\xi)$ whence

$$R_n(x_1) = f^{(n+1)}(\xi) \frac{(x_1 - x_0)^{n+1}}{(n+1)!} \quad \text{where } \xi \in (x_0, x_1)$$

Since x_1 is an arbitrary point the theorem has been proved.

The expression for the remainder found in this theorem makes it possible to rewrite equality (*) as

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots \\ \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1} \quad (**)$$

¹ Let the reader take a concrete value of n , say, $n=4$, and carry out all the calculations in full. This will clarify the argument, concerning an arbitrary value of n , given in the proof.

This is *Taylor's formula with the n th remainder in Lagrange's form*. The only distinction between the last term of this formula (the remainder) and the general term of Taylor's series is that in this term the value of the corresponding derivative is taken not at the point x_0 but at an intermediate point ξ lying between x_0 and x .

In the derivation of Taylor's formula (**) we only supposed that the function $f(x)$ possessed the derivatives up to the order $n+1$ inclusive where n is a given number. The derivatives of order higher than $n+1$ do not necessarily exist. Thus, Taylor's formula can be written for a function differentiable several times even when the whole Taylor series (involving infinitely many derivatives) does not exist.

Let us indicate some particular cases of this formula.

(1) Let $n=0$. Then we obtain Lagrange's formula

$$f(x) = f(x_0) + f'(\xi)(x - x_0)$$

which thus is a special case of Taylor's formula.

(2) Let $n=1$. Then

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2!}(x - x_0)^2$$

If we discard the remainder in this formula we arrive at the approximation to the value of the function $f(x)$ at the point x obtained in Sec. 53 (where the first differential was applied to approximate calculations):

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

In this case the function $f(x)$ is replaced by a linear function.

Lagrange's form of the remainder $R_n(x)$ cannot be used for the exact computation of the value of $R_n(x)$ since the exact location of the point ξ at which the $n+1$ derivative is taken is *unknown*. Therefore formula (**) is usually applied to *estimate* the remainder $R_n(x)$ and the error introduced when it is discarded. Such an estimation is based on the following property:

Let the derivative $f^{(n+1)}(x)$ satisfy the condition

$$|f^{(n+1)}(x)| \leq M_{n+1}$$

(where M_{n+1} is a positive number) in the interval for which Taylor's formula is valid. Then, for any value of x from this interval, the remainder $R_n(x)$ satisfies the inequality

$$|R_n(x)| < M_{n+1} \frac{|x - x_0|^{n+1}}{(n+1)!} \quad (***)$$

Indeed, by the theorem proved above, we have

$$|R_n(x)| = \left| f^{n+1}(\xi) \frac{(x-x_0)^{n+1}}{(n+1)!} \right| = \frac{1}{(n+1)!} |f^{n+1}(\xi)| |x-x_0|^{n+1} \leqslant M_{n+1} \frac{|x-x_0|^{n+1}}{(n+1)!}$$

194. Expanding Functions into Taylor's and Maclaurin's¹ Series. The expansion of a given function $f(x)$ into Taylor's series in the vicinity of a point x_0 reduces to the following two stages:

(1) Under the assumption that the function $f(x)$ is infinitely differentiable the values of $f(x)$ and of its derivatives at the point x_0 are computed and the Taylor series for the function $f(x)$ is formed.

(2) The interval in which the constructed Taylor series converges to the function $f(x)$ is determined, that is the range of the values of x for which the remainder $R_n(x)$ in Taylor's formula tends to zero, as $n \rightarrow \infty$, is found. For this purpose the following theorem is often of use.

Theorem. If the absolute values of all the derivatives of the function $f(x)$ have a common upper bound for an interval containing the point x_0 the function $f(x)$ can be expanded into Taylor's series in this interval.

Proof. By the hypothesis,

$$|f^{(n+1)}(x)| \leqslant M$$

for all the points of the interval in question where M is a constant independent of n . We must prove that the remainder $R_n(x)$ tends to zero as $n \rightarrow \infty$ for all the points of this interval.

By virtue of inequality (***) of Sec. 193, we have

$$|R_n(x)| \leqslant M \frac{|x-x_0|^{n+1}}{(n+1)!}$$

Since the ratio $\frac{|x-x_0|^n}{n!}$ tends to zero for $n \rightarrow \infty$ (see Sec. 189, Example 1) it follows that

$$\lim_{n \rightarrow \infty} R_n(x) = 0$$

for all the points x of the interval, which is what we had to prove.

The expansion of a function in powers of x is most frequently used. In this case we put $x_0=0$ in Taylor's series to obtain the series

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n + \dots$$

A series of this type which is a particular case of Taylor's series

¹ Maclaurin, C. (1698-1746), a Scotch mathematician.

is traditionally called *Maclaurin's series* although this is historically incorrect.

Let us consider the expansions of elementary functions in power series.

I. The exponential function e^x . Let us expand the function

$$f(x) = e^x$$

into Maclaurin's series. All the derivatives of e^x coincide with e^x and take on the value 1 at the point $x=0$. Therefore Maclaurin's series takes the form

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots, \text{ and thus}$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + R_n(x)$$

Let us consider an interval $[-N, N]$ where N is an arbitrary fixed number. For all the values of x from this interval we have $e^x < e^N = M$. Consequently, the absolute values of the derivatives of the function $f(x) = e^x$ have the common upper bound $M = e^N$, and the theorem proved at the beginning of this section implies that $\lim_{n \rightarrow \infty} R_n(x) = 0$. The number N being chosen quite arbitrarily,

we conclude that *the function e^x can be expanded into Maclaurin's series for all the values of x , i.e. throughout the Ox axis.*

Thus,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots \quad (-\infty < x < \infty) \quad (1)$$

The fact that the radius of convergence of series (1) is equal to infinity was proved earlier (see Sec. 189, Example 1). The result established here additionally states that the sum of this series is equal to e^x for any x .

In particular, for $x=1$ we obtain the numerical series $\sum_{n=0}^{\infty} \frac{1}{n!}$ convergent to the number e :

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} + \dots$$

II. Trigonometric functions $\sin x$ and $\cos x$. Let us expand the function $\sin x$ into Maclaurin's series. To this end, we consecutively find the values of its derivatives at the point $x=0$:

$$\begin{aligned} f(0) &= (\sin x)_{x=0} = 0, \quad f'(0) = (\cos x)_{x=0} = 1 \\ f''(0) &= (-\sin x)_{x=0} = 0 \\ f'''(0) &= (-\cos x)_{x=0} = -1, \quad f^{IV}(0) = (\sin x)_{x=0} = 0, \text{ etc.} \end{aligned}$$

We see that the values of the derivatives form the periodic sequence

$$0, 1, 0, -1, 0, 1, 0, -1, \dots$$

Any derivative of the function $\sin x$ (it is equal to $\pm \cos x$ or to $\pm \sin x$) does not exceed unity in its absolute value. Consequently, the Maclaurin series of the function $\sin x$ converges to this function throughout the x -axis.

Thus

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} + \dots \quad (2)$$

$$(-\infty < x < \infty)$$

The expansion of the function $\cos x$ into Maclaurin's series is found quite analogously (this series also converges to $\cos x$ throughout Ox):

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + \dots \quad (3)$$

$$(-\infty < x < \infty)$$

Note that the Maclaurin series of the *odd* function $\sin x$ only involves odd powers of x and that of the *even* function $\cos x$ only even powers of x .

The possibility of expanding the functions e^x , $\sin x$ and $\cos x$ into Maclaurin's series has been easily proved since all the derivatives of these functions can readily be computed. They all turn out to be bounded within any fixed interval of the axis Ox , and therefore the series obtained are convergent to the corresponding functions throughout Ox .

When the remainder is of a more complex structure and it is difficult to investigate its behaviour (in particular, when the theorem proved above is inapplicable) the convergence of Taylor's series to the given function can be tested in the following way. We first form the Taylor series of the given function $f(x)$ and determine its interval of convergence and then try to prove that the remainder in Taylor's formula tends to zero for the values of x belonging to the interval of convergence. There is no need in investigating the remainder for the values of x lying outside the interval of convergence since it cannot tend to zero for these values.

Usually we deal with functions $f(x)$ whose Taylor series converge to them throughout the interval of convergence. An example of a function for which this is not the case was discussed on p. 694.

III. The binomial series. Let us expand into Maclaurin's series the function

$$f(x) = (1+x)^m$$

where m is an arbitrary real number¹. In this case we have

$$f'(x) = m(1+x)^{m-1}, \quad f''(x) = m(m-1)(1+x)^{m-2}, \quad \dots, \quad f^{(n)}(x) = m(m-1)\dots(m-n+1)(1+x)^{m-n}, \dots$$

Therefore, $f(0) = 1$, $f'(0) = m$, $f''(0) = m(m-1)$, $\dots$, $f^{(n)}(0) = m(m-1)\dots(m-n+1)$. Hence, Maclaurin's series is written as

$$1 + mx + \frac{m(m-1)}{2!}x^2 + \dots + \frac{m(m-1)\dots(m-n+1)}{n!}x^n + \dots$$

The latter expression is called the *binomial series*. Let us begin with determining the domain of convergence of this series. To this end we find the limit of the absolute value of the ratio of its two successive terms:

$$\begin{aligned} \lim_{n \rightarrow \infty} \left| \frac{m(m-1)\dots(m-n+1)(m-n)}{(n+1)!} x^{n+1} \left(\frac{m(m-1)\dots(m-n+1)}{n!} x^n \right)^{-1} \right| &= \\ &= |x| \lim_{n \rightarrow \infty} \left| \frac{m-n}{n+1} \right| = |x| \end{aligned}$$

Thus, according to D'Alembert's test, *the series is convergent for $|x| < 1$ and divergent for $|x| > 1$.*

Now we pass to the investigation of the remainder confining ourselves to the case $0 < x < 1$. For this interval and for all $n > m-1$ we have

$$(1+x)^{m-n-1} = \frac{1}{(1+x)^{n-(m-1)}} < 1$$

and, consequently,

$$\begin{aligned} |f^{(n+1)}(x)| &= |m(m-1)\dots(m-n)(1+x)^{m-n-1}| < \\ &< |m(m-1)\dots(m-n)| \end{aligned}$$

The theorem proved at the beginning of the present section is inapplicable here since the upper bound we have found is *dependent on n* . Let us use inequality (***) established in Sec. 193:

$$|R_n(x)| < \left| \frac{m(m-1)\dots(m-n)}{(n+1)!} x^{n+1} \right|$$

The right member of this inequality is equal to the absolute value of the $(n+1)$ th term of the power series whose convergence for $|x| < 1$ has just been proved. Hence, $\lim_{n \rightarrow \infty} R_n(x) = 0$ for $0 < x < 1$.

The corresponding proof for the interval $(-1, 0)$ involves some more sophisticated considerations which we do not present here.

¹ We take the expansion of $(1+x)^m$ but not of the power function x^m since if m is not a positive integer or zero then, beginning with a certain order, the derivatives of x^m do not exist at the point $x=0$.

Thus, the binomial series is convergent to the function $(1+x)^m$ in the interval $(-1, 1)$:

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \dots \\ \dots + \frac{m(m-1)\dots(m-n+1)}{n!}x^n + \dots \quad (-1 < x < 1) \quad (4)$$

If m is a positive integer then of course the series on the right-hand side of (4) contains only $m+1$ nonzero terms and thus turns into Newton's binomial formula. Here we do not investigate the convergence of series (4) at the end points of the interval $(-1, 1)$ and only mention that if $m > 0$ the series is convergent to the function $(1+x)^m$ throughout the closed interval $[-1, 1]$.

Let us write down the binomial series corresponding to the exponents $m = -1$, $m = \frac{1}{2}$ and $m = -\frac{1}{2}$ which are frequently used (in the parentheses we indicate the intervals within which these expansions are valid):

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-1)^n x^n + \dots \quad (-1 < x < 1);$$

(the series on the right is the geometric series with ratio $-x$);

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{2 \cdot 4}x^2 + \frac{1 \cdot 3}{2 \cdot 4 \cdot 6}x^3 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8}x^4 + \dots \\ \dots + (-1)^{n-1} \frac{1 \cdot 3 \cdot 5 \dots (2n-3)}{2 \cdot 4 \cdot 6 \cdot 8 \dots 2n} x^n + \dots \quad (-1 \leq x \leq 1)$$

$$\frac{1}{\sqrt{1+x}} = 1 - \frac{1}{2}x + \frac{1 \cdot 3}{2 \cdot 4}x^2 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^3 + \frac{1 \cdot 3 \cdot 5 \cdot 7}{2 \cdot 4 \cdot 6 \cdot 8}x^4 - \dots \\ \dots + (-1)^n \frac{1 \cdot 3 \cdot 5 \dots (2n-1)}{2 \cdot 4 \cdot 6 \cdot 8 \dots 2n} x^n + \dots \quad (-1 < x \leq 1)$$

There are various functions whose expansions into Taylor's or Maclaurin's series can be found by using the expansions already known and the properties of power series. Examples below demonstrate these techniques.

IV. The functions $\ln(1+x)$ and $\arctan x$. To expand the function $f(x) = \ln(1+x)$ into Maclaurin's series¹ we make use of the formula for the sum of the geometric series

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-1)^n x^n + \dots \\ (-1 < x < 1) \quad (*)$$

¹ The function $y = \ln x$ cannot be expanded into Maclaurin's series at the point $x=0$ since it is discontinuous at that point.

Let us apply the theorem on the termwise integration of power series. Since

$$\int_0^x \frac{dx}{1+x} = \ln(1+x) \Big|_0^x = \ln(1+x)$$

the integration of series (*) from 0 to x results in

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^{n-1} \frac{x^n}{n} + \dots$$

$$(-1 < x < 1) \quad (5)$$

Series (5) diverges for $|x| > 1$ and also at the point $x = -1$, and therefore does not represent the function $\ln(1+x)$ for these values of x . The series is convergent at the point $x = 1$, and the additional remark to the first property of power series made in Sec. 190 indicates that its sum is continuous (from the left) at the end point $x = 1$. Therefore expansion (5) remains valid for $x = 1$:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + (-1)^{n-1} \frac{1}{n} + \dots$$

The series on the right converges very slowly and is therefore inapplicable in practical computation of $\ln 2$. Using the estimation given in Leibniz' test for alternating series we readily find that in order to compute $\ln 2$ to an accuracy of 0.00001 with the aid of this series we should take 100,000 terms, which is senseless from the point of view of practice! Later on (see Sec. 195) we shall show that an appropriate transformation of this series (which *accelerates* its convergence) makes it possible to achieve the desired accuracy in the computation of $\ln 2$ with an essential reduction of the amount of calculations.

The expansion of the function $\arctan x$ into Maclaurin's series is found completely analogously. To this end we substitute x^2 for x in the sum of geometric series (*) to obtain

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots + (-1)^n x^{2n} + \dots \quad (-1 < x < 1) \quad (**)$$

On integrating series (**) from 0 to x ($|x| < 1$) we receive

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + \dots$$

$$(-1 < x < 1) \quad (6)$$

This series remains convergent at the end points of the interval of convergence, and therefore expansion (6) is also valid for $x = \pm 1$. Substituting the value $x = 1$ into this series we obtain the representation of the number π in the form of the numerical

series

$$\arctan 1 = \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$$

It is advisable to memorize formulas (1)-(6) and the corresponding intervals of convergence of the expansions since there are many other functions whose expansions into power series can be found by using these formulas.

Let us consider several examples.

Examples. (1) Let us expand into Maclaurin's series the hyperbolic functions $\cosh x$ and $\sinh x$. To this end, we take series (1) for the function e^x and replace x by $-x$; this yields

$$e^{-x} = 1 - x + \frac{x^2}{2!} - \frac{x^3}{3!} + \dots + (-1)^n \frac{x^n}{n!} + \dots$$

Now, using the addition and subtraction rules for series we find the desired expansions valid for all the values of the argument x :

$$\cosh x = \frac{e^x + e^{-x}}{2} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + \frac{x^{2n}}{(2n)!} + \dots$$

and

$$\sinh x = \frac{e^x - e^{-x}}{2} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + \frac{x^{2n-1}}{(2n-1)!} + \dots$$

(2) Let us expand into Maclaurin's series the function

$$y = e^{-x^2}$$

We again take the expansion of the function $y = e^x$ and substitute $-x^2$ for x into it, which results in

$$e^{-x^2} = 1 - x^2 + \frac{x^4}{2!} - \frac{x^6}{3!} + \dots$$

This expansion is valid throughout the x -axis.

(3) Let us take the function

$$y = \frac{x^3}{\sqrt{1-x^2}}$$

and expand it into Maclaurin's series. In this case we make use of the binomial series for $m = -\frac{1}{2}$, that is, of the expansion of the function $\frac{1}{\sqrt{1+x}}$ (see the examples above); on replacing x by $-x^2$ in this series we receive

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{1}{2}x^2 + \frac{1 \cdot 3}{2 \cdot 4}x^4 + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^6 + \dots \quad (***)$$

Let us apply the theorem on the termwise integration of power series. Since

$$\int_0^x \frac{dx}{1+x} = \ln(1+x) \Big|_0^x = \ln(1+x)$$

the integration of series (*) from 0 to x results in

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^{n-1} \frac{x^n}{n} + \dots$$

$$(-1 < x < 1) \quad (5)$$

Series (5) diverges for $|x| > 1$ and also at the point $x = -1$, and therefore does not represent the function $\ln(1+x)$ for these values of x . The series is convergent at the point $x = 1$, and the additional remark to the first property of power series made in Sec. 190 indicates that its sum is continuous (from the left) at the end point $x = 1$. Therefore expansion (5) remains valid for $x = 1$:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + (-1)^{n-1} \frac{1}{n} + \dots$$

The series on the right converges very slowly and is therefore inapplicable in practical computation of $\ln 2$. Using the estimation given in Leibniz' test for alternating series we readily find that in order to compute $\ln 2$ to an accuracy of 0.00001 with the aid of this series we should take 100,000 terms, which is senseless from the point of view of practice! Later on (see Sec. 195) we shall show that an appropriate transformation of this series (which *accelerates* its convergence) makes it possible to achieve the desired accuracy in the computation of $\ln 2$ with an essential reduction of the amount of calculations.

The expansion of the function $\arctan x$ into Maclaurin's series is found completely analogously. To this end we substitute x^2 for x in the sum of geometric series (*) to obtain

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \dots + (-1)^n x^{2n} + \dots \quad (-1 < x < 1) \quad (**)$$

On integrating series (**) from 0 to x ($|x| < 1$) we receive

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \dots + (-1)^{n-1} \frac{x^{2n-1}}{2n-1} + \dots$$

$$(-1 < x < 1) \quad (6)$$

This series remains convergent at the end points of the interval of convergence, and therefore expansion (6) is also valid for $x = \pm 1$. Substituting the value $x = 1$ into this series we obtain the representation of the number π in the form of the numerical

tive we can try to construct a dominant series with a known sum for the remainder of the given series (for this purpose we most often use a geometric series).

In practice the estimation of the terms of the remainder series is often more convenient than the investigation of the general expression for the remainder term (i.e. for the sum of the remainder series) since the latter requires the knowledge of the derivative of the corresponding order of the given function in the whole interval in question. A power series expansion of a function can sometimes be found by combining series already investigated without computing the derivatives. These techniques are illustrated by the examples below.

Examples. (1) We have

$$e^x \approx 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}$$

Using this approximate formula for an interval $[0, M]$ where M is an arbitrary fixed number and taking into account that the greatest value of the derivative $f^{(n+1)}(x) = e^x$ attained on this interval is equal to e^M we can apply the theorem on estimating the remainder to obtain

$$R_n(x) < e^M \frac{x^{n+1}}{(n+1)!}$$

This estimation is inconvenient since it involves the number e^M . Moreover, it may prove excessive since e^M is the greatest value of the derivative $f^{(n+1)}(x) = e^x$ for the whole interval $[0, M]$.

For $M=1$ we obtain¹

$$R_n(x) < e \frac{x^{n+1}}{(n+1)!} < 3 \frac{x^{n+1}}{(n+1)!}$$

For example, let us estimate the number of terms needed for computing the number e to within 0.00001. Here we put $x=1$ and obtain

$$\frac{3}{(n+1)!} < 10^{-5}$$

It is readily checked that this inequality holds for $n=8$: $\frac{3}{9!} = \frac{3}{362880} < 10^{-5}$. Taking into account the errors of the arithmetical operations we conclude that to compute the number e with the desired accuracy and to avoid the accumulation of rounding errors we must evaluate all the terms in the sum

$$e \approx 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{8!} \approx 2.71828$$

¹ For the sake of simplicity we intentionally increase the right-hand side.

The required expansion of the given function is now obtained by multiplying both sides of equality (***) by x^3 :

$$\frac{x^3}{\sqrt{1-x^2}} = x^3 + \frac{1}{2}x^5 + \frac{1 \cdot 3}{2 \cdot 4}x^7 + \dots$$

(4) Let us find the expansion in Maclaurin's series of the function

$$y = \arcsin x$$

Since

$$\int_0^x \frac{dx}{\sqrt{1-x^2}} = \arcsin x \Big|_0^x = \arcsin x$$

the expansion of $\arcsin x$ is readily found by integrating series (***):

$$\arcsin x = x + \frac{1}{2} \frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \frac{x^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \frac{x^7}{7} + \dots$$

The latter expansion is valid for the closed interval $[-1, 1]$.

(5) Let us expand into Maclaurin's series the function

$$y = e^x \sin x$$

The Maclaurin series of the functions e^x and $\sin x$ being absolutely convergent, the sought-for expansion can be obtained by the multiplication of these series according to the rule stated at the end of Sec. 186. In this case it is difficult to derive the general expression for the n th term of the resultant series and we only compute the first several terms:

$$e^x \sin x = x + x^2 + \frac{x^3}{3} - \frac{x^6}{30} + \dots$$

§ 5. Some Applications of Taylor's Series

195. Approximate Computation of Values of Functions. Suppose that the values of a given function $f(x)$ and of its successive derivatives at a point x_0 are known and that we managed to prove that the function $f(x)$ can be expanded into Taylor's series in a neighbourhood of the point x_0 . Then the precise value of the function $f(x)$ in the vicinity of the point x_0 can be found by means of Taylor's series and an approximation to this value can be obtained if we take a partial sum of the series. The error of such an approximation can be estimated on the basis of the general theorem on the estimation of the remainder stated in Sec. 194 or, if this theorem proves inapplicable, by investigating directly the expression of the remainder. For instance, if the resultant series is alternating we can use Leibniz' theorem; if the series is posi-

tive we can try to construct a dominant series with a known sum for the remainder of the given series (for this purpose we most often use a geometric series).

In practice the estimation of the terms of the remainder series is often more convenient than the investigation of the general expression for the remainder term (i.e. for the sum of the remainder series) since the latter requires the knowledge of the derivative of the corresponding order of the given function in the whole interval in question. A power series expansion of a function can sometimes be found by combining series already investigated without computing the derivatives. These techniques are illustrated by the examples below.

Examples. (1) We have

$$e^x \approx 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!}$$

Using this approximate formula for an interval $[0, M]$ where M is an arbitrary fixed number and taking into account that the greatest value of the derivative $f^{(n+1)}(x) = e^x$ attained on this interval is equal to e^M we can apply the theorem on estimating the remainder to obtain

$$R_n(x) < e^M \frac{x^{n+1}}{(n+1)!}$$

This estimation is inconvenient since it involves the number e^M . Moreover, it may prove excessive since e^M is the greatest value of the derivative $f^{(n+1)}(x) = e^x$ for the whole interval $[0, M]$.

For $M=1$ we obtain¹

$$R_n(x) < e \frac{x^{n+1}}{(n+1)!} < 3 \frac{x^{n+1}}{(n+1)!}$$

For example, let us estimate the number of terms needed for computing the number e to within 0.00001. Here we put $x=1$ and obtain

$$\frac{3}{(n+1)!} < 10^{-5}$$

It is readily checked that this inequality holds for $n=8$: $\frac{3}{9!} = \frac{3}{362880} < 10^{-5}$. Taking into account the errors of the arithmetical operations we conclude that to compute the number e with the desired accuracy and to avoid the accumulation of rounding errors we must evaluate all the terms in the sum

$$e \approx 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{8!} \approx 2.71828$$

¹ For the sake of simplicity we intentionally increase the right-hand side.

with an accuracy of 10^{-6} . This results in an approximation to the number e correct to six decimal places.

Let the reader verify that in order to compute $\sqrt[n]{e} = e^{\frac{1}{n}}$ with the same accuracy it is sufficient to take $n=6$, that is to evaluate seven summands. Generally, the smaller the index of the power of e , the smaller the number of terms in the approximate formula providing the required accuracy.

Now we shall show how the whole Taylor series of e^x can be used for estimating the error $R_n(1)$:

$$\begin{aligned} R_n(1) &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \dots = \\ &= \frac{1}{(n+1)!} \left[1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \dots \right] < \\ &< \frac{1}{(n+1)!} \left[1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \dots \right] = \frac{1}{(n+1)!} \frac{1}{1 - \frac{1}{n+1}} = \frac{1}{n!n} \end{aligned}$$

We have thus found the upper bound $\frac{1}{n!n}$ for the absolute error which is almost three times as accurate as the value $\frac{3}{(n+1)!}$ obtained earlier.

Let us compute approximations to the values of the functions $\sin x$ and $\cos x$ with the aid of their power series expansions. In this case the approximate formulas are extremely precise and the error is easily estimated by means of Leibniz' theorem. Taking partial sums of the Maclaurin series of the function $\sin x$ and assuming that $x > 0$ we obtain, in succession, the approximate formulas

$$\begin{aligned} \sin x &\approx x, \quad |R_1(x)| < \frac{x^3}{6} \\ \sin x &\approx x - \frac{x^3}{6}, \quad |R_3(x)| < \frac{x^5}{120} \\ \sin x &\approx x - \frac{x^3}{6} + \frac{x^5}{120}, \quad |R_5(x)| < \frac{x^7}{5040} \end{aligned}$$

The first and the third formulas provide major approximations to the values of $\sin x$ and the second gives a minor approximation. Let the reader check that to compute the values of $\sin x$ with an accuracy of 0.0001 we should use the first formula for the interval $0 < x < 0.08$, the second formula for the interval $0.08 < x < 0.4$ and the third for $0.4 < x < 0.9$.¹

¹ Accordingly, in the sexagesimal measure, these intervals are $(0^\circ, 4^\circ 30')$, $(4^\circ 30', 23^\circ)$ and $(23^\circ, 52^\circ)$ (the boundaries of the intervals are computed approximately).

Fig. 249 illustrates the graph of the function $y = \sin x$ and the graphs of the approximating Taylor polynomials

$$y = x, \quad y = x - \frac{x^3}{6} \quad \text{and} \quad y = x - \frac{x^3}{6} + \frac{x^5}{120}$$

in the vicinity of the point $x = 0$.

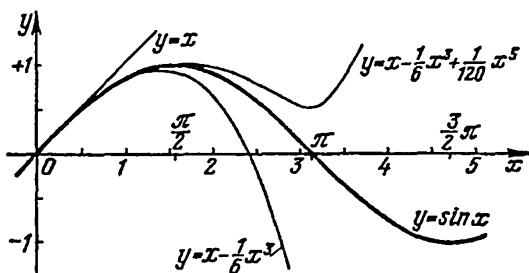


Fig. 249

For the function $y = \cos x$ we similarly obtain the approximate formulas

$$\cos x \approx 1, \quad |R_0(x)| < \frac{x^2}{2}$$

$$\cos x \approx 1 - \frac{x^2}{2}, \quad |R_2(x)| < \frac{x^4}{24}$$

$$\cos x \approx 1 - \frac{x^2}{2} + \frac{x^4}{24}, \quad |R_4(x)| < \frac{x^6}{720}$$

The first and the third formulas give major approximations and the second gives a minor approximation to the values of $\cos x$.

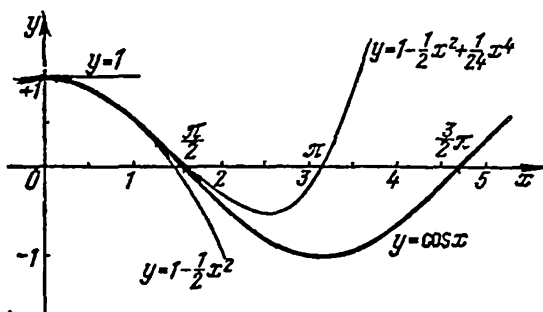


Fig. 250

In Fig. 250 we see the graphs of the function $y = \cos x$ and of the approximating Taylor polynomials

$$y = 1, \quad y = 1 - \frac{x^2}{2} \quad \text{and} \quad y = 1 - \frac{x^2}{2} + \frac{x^4}{24}$$

in the neighbourhood of the origin.

(3) Let us take the power series for $\ln(1+x)$ derived earlier:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad (-1 < x \leq 1)$$

The rate of decrease of its coefficients as $n \rightarrow \infty$ is much lower than that of the coefficients of the Maclaurin series of the functions e^x , $\sin x$ and $\cos x$, and therefore the convergence of the series is much slower. For instance, as was mentioned, Leibniz' test for alternating series shows that to compute $\ln 2$ with the aid of this series to an accuracy of 0.00001 it is necessary to take not less than 100,000 terms of the series, which is not realizable practically.

Let us show how the convergence of this series can be accelerated. To this end we replace x by $-x$ and write down the series for $\ln(1-x)$. On subtracting the latter series from the Maclaurin series of $\ln(1+x)$ we obtain

$$\ln \frac{1+x}{1-x} = 2 \left(x + \frac{1}{3} x^3 + \frac{1}{5} x^5 + \dots \right) \quad (-1 < x < 1)$$

This formula enables us to compute the logarithm of any positive number. Indeed, if x ranges within the interval of convergence of the series (i.e. $x \in (-1, 1)$) the value of the continuous function $\frac{1+x}{1-x}$ runs through the whole interval $(0, \infty)$. Let us apply this series to compute $\ln 2$. If $\frac{1+x}{1-x} = 2$ then $x = \frac{1}{3}$. The n th partial sum of the series gives us the approximation

$$\ln 2 \approx 2 \left(\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3^3} + \dots + \frac{1}{2n+1} \cdot \frac{1}{3^{2n+1}} \right)$$

To estimate the error we use the inequality

$$\begin{aligned} 2 \left(\frac{1}{2n+3} \cdot \frac{1}{3^{2n+3}} + \frac{1}{2n+5} \cdot \frac{1}{3^{2n+5}} + \dots \right) &< \\ &< \frac{2}{(2n+3) \cdot 3^{2n+3}} \left(1 + \frac{1}{3^2} + \frac{1}{3^4} + \dots \right) = \\ &= \frac{2 \cdot 9}{(2n+3) \cdot 3^{2n+3} \cdot 8} = \frac{1}{4(2n+3) \cdot 3^{2n+1}} \end{aligned}$$

Let us determine the value of n for which the error does not exceed 0.00001. We must have the inequality

$$4(2n+3) \cdot 3^{2n+1} > 10^5$$

which is fulfilled beginning with $n=4$. Thus,

$$\ln 2 \approx 2 \left(\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3^3} + \frac{1}{5} \cdot \frac{1}{3^5} + \frac{1}{7} \cdot \frac{1}{3^7} + \frac{1}{9} \cdot \frac{1}{3^9} \right) \approx 0.693144$$

We see that it is sufficient to take only five terms of the new series (instead of 100,000 terms of the original series) to compute the desired value $\ln 2$ with the same accuracy.

To apply this series to the computation of $\ln(N+1)$ proceeding from the value of $\ln N$ already found where N is a positive integer we should put $x = \frac{1}{2N+1}$ in the series, which leads to the formula

$$\ln \frac{N+1}{N} = \ln(N+1) - \ln N = 2 \left(\frac{1}{2N+1} + \frac{1}{3} \cdot \frac{1}{(2N+1)^3} + \dots \right)$$

serving for practical computation with an arbitrary accuracy of the logarithms of positive integers.

196. Integrating Functions. Application of Taylor's Series to Differential Equations.

I. Integrating functions. Suppose that it is required to find the integral

$$F(x) = \int_a^x f(x) dx$$

where the integrand $f(x)$ is a function whose Taylor series is known, the limits of integration lying within the interval of convergence of the series. Then it is allowable to integrate the series termwise. This results in the Taylor series of the function $F(x)$ whose radius of convergence coincides with that of the series for

the original function $f(x)$. If the integral $\int_a^x f(x) dx$ is expressible as an elementary function $F(x)$ we thus obtain its expansion into Taylor's series. For instance, this was the case in the examples where we found the expansions of the functions $\ln(1+x)$, $\arctan x$

and $\arcsin x$. If the integral $\int_a^x f(x) dx$ is inexpressible in terms of elementary functions (see Sec. 85) the new series expresses the *nonelementary* function $F(x)$ in terms of basic elementary functions (i.e. in terms of power functions) but with the aid of an *infinite* number of elementary operations.

It should be noted that the application of the theorem on the estimation of the definite integral makes it possible to estimate

the remainder term of the series for the integral $F(x) = \int_a^x f(x) dx$ if an estimation for the remainder of the Taylor series of the integrand function $f(x)$ is known.

(3) Let us take the power series for $\ln(1+x)$ derived earlier:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots \quad (-1 < x \leq 1)$$

The rate of decrease of its coefficients as $n \rightarrow \infty$ is much lower than that of the coefficients of the Maclaurin series of the functions e^x , $\sin x$ and $\cos x$, and therefore the convergence of the series is much slower. For instance, as was mentioned, Leibniz' test for alternating series shows that to compute $\ln 2$ with the aid of this series to an accuracy of 0.00001 it is necessary to take not less than 100,000 terms of the series, which is not realizable practically.

Let us show how the convergence of this series can be accelerated. To this end we replace x by $-x$ and write down the series for $\ln(1-x)$. On subtracting the latter series from the Maclaurin series of $\ln(1+x)$ we obtain

$$\ln \frac{1+x}{1-x} = 2 \left(x + \frac{1}{3} x^3 + \frac{1}{5} x^5 + \dots \right) \quad (-1 < x < 1)$$

This formula enables us to compute the logarithm of any positive number. Indeed, if x ranges within the interval of convergence of the series (i.e. $x \in (-1, 1)$) the value of the continuous function $\frac{1+x}{1-x}$ runs through the whole interval $(0, \infty)$. Let us apply this series to compute $\ln 2$. If $\frac{1+x}{1-x} = 2$ then $x = \frac{1}{3}$. The n th partial sum of the series gives us the approximation

$$\ln 2 \approx 2 \left(\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3^3} + \dots + \frac{1}{2n+1} \cdot \frac{1}{3^{2n+1}} \right)$$

To estimate the error we use the inequality

$$\begin{aligned} 2 \left(\frac{1}{2n+3} \cdot \frac{1}{3^{2n+3}} + \frac{1}{2n+5} \cdot \frac{1}{3^{2n+5}} + \dots \right) &< \\ &< \frac{2}{(2n+3) \cdot 3^{2n+3}} \left(1 + \frac{1}{3^2} + \frac{1}{3^4} + \dots \right) = \\ &= \frac{2 \cdot 9}{(2n+3) \cdot 3^{2n+3} \cdot 8} = \frac{1}{4(2n+3) \cdot 3^{2n+1}} \end{aligned}$$

Let us determine the value of n for which the error does not exceed 0.00001. We must have the inequality

$$4(2n+3) \cdot 3^{2n+1} > 10^5$$

which is fulfilled beginning with $n=4$. Thus,

$$\ln 2 \approx 2 \left(\frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3^3} + \frac{1}{5} \cdot \frac{1}{3^5} + \frac{1}{7} \cdot \frac{1}{3^7} + \frac{1}{9} \cdot \frac{1}{3^9} \right) \approx 0.693144$$

We see that it is sufficient to take only five terms of the new series (instead of 100,000 terms of the original series) to compute the desired value $\ln 2$ with the same accuracy.

To apply this series to the computation of $\ln(N+1)$ proceeding from the value of $\ln N$ already found where N is a positive integer we should put $x = \frac{1}{2N+1}$ in the series, which leads to the formula

$$\ln \frac{N+1}{N} = \ln(N+1) - \ln N = 2 \left(\frac{1}{2N+1} + \frac{1}{3} \cdot \frac{1}{(2N+1)^3} + \dots \right)$$

serving for practical computation with an arbitrary accuracy of the logarithms of positive integers.

196. Integrating Functions. Application of Taylor's Series to Differential Equations.

I. Integrating functions. Suppose that it is required to find the integral

$$F(x) = \int_a^x f(x) dx$$

where the integrand $f(x)$ is a function whose Taylor series is known, the limits of integration lying within the interval of convergence of the series. Then it is allowable to integrate the series termwise. This results in the Taylor series of the function $F(x)$ whose radius of convergence coincides with that of the series for

the original function $f(x)$. If the integral $\int_a^x f(x) dx$ is expressible

as an elementary function $F(x)$ we thus obtain its expansion into Taylor's series. For instance, this was the case in the examples where we found the expansions of the functions $\ln(1+x)$, $\arctan x$

and $\arcsin x$. If the integral $\int_a^x f(x) dx$ is inexpressible in terms of elementary functions (see Sec. 85) the new series expresses the *nonelementary* function $F(x)$ in terms of basic elementary functions (i.e. in terms of power functions) but with the aid of an *infinite* number of elementary operations.

It should be noted that the application of the theorem on the estimation of the definite integral makes it possible to estimate

the remainder term of the series for the integral $F(x) = \int_a^x f(x) dx$

if an estimation for the remainder of the Taylor series of the integrand function $f(x)$ is known.

Examples. (1) Consider the integral

$$\int_0^x \frac{\sin x}{x} dx$$

Dividing the Maclaurin series of $\sin x$ by x we obtain

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots$$

This series, like the series for $\sin x$, is convergent throughout the Ox -axis. The integration yields

$$\int_0^x \frac{\sin x}{x} dx = x - \frac{x^3}{3! \cdot 3} + \frac{x^5}{5! \cdot 5} - \dots$$

There exists no *elementary function* to which the latter series converges; it specifies a new *nonelementary function* by means of an infinite number of arithmetical operations. By the way, this function is used in some divisions of theoretical physics, being

denoted as $\text{Si}(x) = \int_0^x \frac{\sin x}{x} dx$ and called the *sine integral*. There are extensive tables of values of the function $\text{Si}(x)$ compiled with the help of the series obtained here.

(2) In the theory of probability an important role is played by the function

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$$

called *Laplace's function* or the *probability integral*.

This integral cannot be computed in the finite form since $\int e^{-\frac{x^2}{2}} dx$ is inexpressible in terms of elementary functions. Let us expand the integrand into a power series; to this end, we substitute $-\frac{x^2}{2}$ for x into the expansion of e^x :

$$e^{-\frac{x^2}{2}} = 1 - \frac{x^2}{2} + \frac{x^4}{2^2 \cdot 2!} - \frac{x^6}{2^3 \cdot 3!} + \dots$$

The term-by-term integration now yields the representation of the function $\Phi(x)$ in the form of the infinite series

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \left(x - \frac{x^3}{3 \cdot 2} + \frac{x^5}{5 \cdot 2^2 \cdot 2!} - \frac{x^7}{7 \cdot 2^3 \cdot 3!} + \dots \right)$$

convergent throughout the Ox -axis. It converges particularly rapidly for $|x| \leq 1$.

(3) Let us compute the integral

$$\int_0^{\frac{1}{2}} \frac{dx}{1+x^4}$$

The indefinite integral $\int \frac{dx}{1+x^4}$ is expressible in elementary functions, and it may seem that there is no need in expanding the integrand into a power series. But the matter is that the expression of this integral is rather complicated and therefore inconvenient for practical calculations:

$$\int \frac{dx}{1+x^4} = \frac{1}{4\sqrt{2}} \ln \frac{x^2 + x\sqrt{2} + 1}{x^2 - x\sqrt{2} + 1} + \frac{\sqrt{2}}{4} \arctan \frac{x\sqrt{2}}{1-x^2} + C$$

At the same time, the series expansion

$$\frac{1}{1+x^4} = 1 - x^4 + x^8 - x^{12} + \dots$$

of the integrand immediately gives us

$$\int_0^{\frac{1}{2}} \frac{dx}{1+x^4} = \frac{1}{2} - \left(\frac{1}{2}\right)^5 \cdot \frac{1}{5} + \left(\frac{1}{2}\right)^9 \cdot \frac{1}{9} - \dots$$

If we confine ourselves to the first two terms of the series (the corresponding error then does not exceed $\left(\frac{1}{2}\right)^9 \cdot \frac{1}{9} \approx 0.0002$) we obtain the approximation

$$\int_0^{\frac{1}{2}} \frac{dx}{1+x^4} \approx 0.4938$$

II. Solving differential equations. We shall consider the application of power series to solving differential equations. Let there be given a differential equation with some initial conditions specifying a particular solution¹. Suppose that this particular solution can be expanded into a power series in the vicinity of the point x_0 for which the initial conditions are set. Such a series arranged in ascending powers of the difference $x - x_0$ is written as

$$y = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots$$

Regarding $a_0, a_1, \dots$ as undetermined coefficients we can differentiate this series as many times as the order of the differential

¹ Power series sometimes also enable us to find the general solution of a differential equation but we do not treat this question here referring the reader to special books, for instance to [12].

Examples. (1) Consider the integral

$$\int_0^x \frac{\sin x}{x} dx$$

Dividing the Maclaurin series of $\sin x$ by x we obtain

$$\frac{\sin x}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \dots$$

This series, like the series for $\sin x$, is convergent throughout the Ox -axis. The integration yields

$$\int_0^x \frac{\sin x}{x} dx = x - \frac{x^3}{3!3} + \frac{x^5}{5!5} - \dots$$

There exists no *elementary function* to which the latter series converges; it specifies a new *nonelementary function* by means of an infinite number of arithmetical operations. By the way, this function is used in some divisions of theoretical physics, being

denoted as $\text{Si}(x) = \int_0^x \frac{\sin x}{x} dx$ and called the *sine integral*. There

are extensive tables of values of the function $\text{Si}(x)$ compiled with the help of the series obtained here.

(2) In the theory of probability an important role is played by the function

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \int_0^x e^{-\frac{x^2}{2}} dx$$

called *Laplace's function* or the *probability integral*.

This integral cannot be computed in the finite form since $\int e^{-\frac{x^2}{2}} dx$ is inexpressible in terms of elementary functions. Let us expand the integrand into a power series; to this end, we substitute $-\frac{x^2}{2}$ for x into the expansion of e^x :

$$e^{-\frac{x^2}{2}} = 1 - \frac{x^2}{2} + \frac{x^4}{2^2 \cdot 2!} - \frac{x^6}{2^3 \cdot 3!} + \dots$$

The term-by-term integration now yields the representation of the function $\Phi(x)$ in the form of the infinite series

$$\Phi(x) = \sqrt{\frac{2}{\pi}} \left(x - \frac{x^3}{3 \cdot 2} + \frac{x^5}{5 \cdot 2^2 \cdot 2!} - \frac{x^7}{7 \cdot 2^3 \cdot 3!} + \dots \right)$$

convergent throughout the Ox -axis. It converges particularly rapidly for $|x| \leq 1$.

(3) Let us compute the integral

$$\int_0^{\frac{1}{2}} \frac{dx}{1+x^4}$$

The indefinite integral $\int \frac{dx}{1+x^4}$ is expressible in elementary functions, and it may seem that there is no need in expanding the integrand into a power series. But the matter is that the expression of this integral is rather complicated and therefore inconvenient for practical calculations:

$$\int \frac{dx}{1+x^4} = \frac{1}{4\sqrt{2}} \ln \frac{x^2 + x\sqrt{2} + 1}{x^2 - x\sqrt{2} + 1} + \frac{\sqrt{2}}{4} \arctan \frac{x\sqrt{2}}{1-x^2} + C$$

At the same time, the series expansion

$$\frac{1}{1+x^4} = 1 - x^4 + x^8 - x^{12} + \dots$$

of the integrand immediately gives us

$$\int_0^{\frac{1}{2}} \frac{dx}{1+x^4} = \frac{1}{2} - \left(\frac{1}{2}\right)^5 \cdot \frac{1}{5} + \left(\frac{1}{2}\right)^9 \cdot \frac{1}{9} - \dots$$

If we confine ourselves to the first two terms of the series (the corresponding error then does not exceed $\left(\frac{1}{2}\right)^9 \cdot \frac{1}{9} \approx 0.0002$) we obtain the approximation

$$\int_0^{\frac{1}{2}} \frac{dx}{1+x^4} \approx 0.4938$$

II. Solving differential equations. We shall consider the application of power series to solving differential equations. Let there be given a differential equation with some initial conditions specifying a particular solution¹. Suppose that this particular solution can be expanded into a power series in the vicinity of the point x_0 for which the initial conditions are set. Such a series arranged in ascending powers of the difference $x - x_0$ is written as

$$y = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \dots$$

Regarding $a_0, a_1, \dots$ as undetermined coefficients we can differentiate this series as many times as the order of the differential

¹ Power series sometimes enable us to find the general solution of a differential equation but we do not treat this question here referring the reader to special books, for instance to [12].

equation. On substituting the series obtained for the unknown function and its derivatives into the differential equation we arrive at an identity from which the coefficients of the series can be found. In this procedure the first coefficients of the series (their number is equal to the order of the differential equation) are determined not from this identity but from the initial conditions. If we manage to prove that the resultant series is convergent it is sure to represent the sought-for solution. In many practical cases there is no need in such a proof. Taking a sufficient number of terms of the series we can obtain an arbitrarily accurate approximation for the solution in the form of a polynomial.

Linear differential equations are particularly convenient for the application of series to finding their solutions.

Let us demonstrate this method by several examples.

Examples. (1) Let us solve the second-order linear differential equation $y'' - xy = 0$ with initial conditions $y|_{x=0} = 0$, $y'|_{x=0} = 1$.

The solution is looked for in the form of a series in powers of x :

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n + \dots$$

The coefficients a_0 and a_1 are found from the initial conditions:

$$a_0 = y|_{x=0} = 0, \quad a_1 = y'|_{x=0} = 1$$

On differentiating the series twice we obtain

$$y'' = 2a_2 + 3 \cdot 2a_3x + \dots + n(n-1)a_nx^{n-2} + \dots$$

The substitution of the series expressing y and y'' into the differential equation results in the identity

$$\begin{aligned} 2a_2 + 3 \cdot 2a_3x + \dots + n(n-1)a_nx^{n-2} + \dots = \\ = a_0x + a_1x^2 + a_2x^3 + \dots + a_{n-3}x^{n-2} + \dots \end{aligned}$$

The comparison of the coefficients in the same powers of x on both sides gives the relations

$$a_2 = 0, \quad a_3 = 0, \quad 4 \cdot 3a_4 = 1, \quad \dots, \quad n(n-1)a_n = a_{n-3}$$

Therefore,

$$\begin{aligned} a_4 = \frac{1}{3 \cdot 4}, \quad a_5 = 0, \quad a_6 = 0, \quad a_7 = \frac{1}{3 \cdot 4 \cdot 6 \cdot 7}, \quad a_8 = 0, \quad a_9 = 0, \\ a_{10} = \frac{1}{3 \cdot 4 \cdot 6 \cdot 7 \cdot 9 \cdot 10} \end{aligned}$$

and, generally,

$$a_{3m-1} = a_{3m} = 0, \quad a_{3m+1} = \frac{1}{3 \cdot 4 \cdot 6 \cdot 7 \dots 3m(3m+1)}$$

Hence,

$$y = x + \frac{1}{3 \cdot 4}x^4 + \frac{1}{3 \cdot 4 \cdot 6 \cdot 7}x^7 + \dots + \frac{1}{3 \cdot 4 \cdot 6 \cdot 7 \dots 3m(3m+1)}x^{3m+1} + \dots$$

Using D'Alembert's test we readily prove that the series obtained is convergent throughout Ox and thus represents the desired solution for all the values of x .

It should be noted that this scheme is applicable to any linear differential equation *irrespective of its order*.

(2) Let us take the equation $y' = xy^2 + 1$ with the initial condition $y|_{x=1} = 0$.

This equation is *nonlinear*, and therefore the substitution of the series

$$y = a_0 + a_1(x-1) + a_2(x-1)^2 + \dots \quad (*)$$

for y into the equation would lead to complicated relations for determining the coefficients. Therefore it is advisable to use another technique. Let us successively differentiate the equation several times regarding y as a function of x :

$$\begin{aligned} y'' &= y^2 + 2xyy' \\ y''' &= 4yy' + 2xy'^2 + 2xyy'' \\ y^{(4)} &= 6y'^2 + 6yy'' + 6xy'y'' + 2xyy''' \end{aligned}$$

Now, putting $x=1$ in the equation itself and in all these equalities and taking into account that $y|_{x=1} = 0$ we find, in succession,

$$y'|_{x=1} = 1, \quad y''|_{x=1} = 0, \quad y'''|_{x=1} = 2, \quad y^{(4)}|_{x=1} = 6$$

Since

$$\begin{aligned} a_0 &= y|_{x=1} = 0, \quad a_1 = y'|_{x=1} = 1, \quad a_2 = \frac{1}{2!} y''|_{x=1} = 0, \\ a_3 &= \frac{1}{3!} y'''|_{x=1} = \frac{1}{3}, \quad a_4 = \frac{1}{4!} y^{(4)}|_{x=1} = \frac{1}{4} \end{aligned}$$

we have

$$y \approx (x-1) + \frac{(x-1)^3}{3} + \frac{(x-1)^4}{4}$$

We did not derive the general formula for the coefficients of series (*) and therefore cannot test it for convergence. In many practical cases a polynomial thus constructed provides a sufficiently accurate approximation to the sought-for solution in a small neighbourhood of the initial point (in our case this is the point $x=1$). The validity of such an approximate solution is usually checked by comparing it with experimental data related to the concrete problem leading to the differential equation under investigation.

§ 6*. Some Further Topics in the Theory of Power Series

197*. Power Series in Complex Argument. Most facts of the theory of real series we have studied are extended almost without changes to series whose terms are *complex* numbers. To consider this generalization we first of all state the definition of the limit of a sequence of complex numbers; the relationship between this definition and the definition of the limit of a real sequence is coherent with that between the definition of the limit of a complex function of a real argument (see Sec. 73) and the definition of the limit of a real function.

Definition. A complex number Z is called *the limit of a sequence of complex numbers* $z_1, z_2, \dots, z_n, \dots$ if, given an arbitrary positive number ε , there is a number N such that the inequality $n > N$ implies the inequality

$$|z_n - Z| < \varepsilon$$

If $z_n = x_n + iy_n$ and $Z = a + ib$ then $|z_n - Z| = \sqrt{(x_n - a)^2 + (y_n - b)^2}$, whence it follows (like in Sec. 73) that the existence of the limit $\lim_{n \rightarrow \infty} z_n = Z$ is equivalent to the existence of the limits of the sequences of real numbers $\lim_{n \rightarrow \infty} x_n = a$ and $\lim_{n \rightarrow \infty} y_n = b$.

This definition allows us to generalize, without any changes, the definition of the convergence of a real series to series with complex terms.

A series

$$w_1 + w_2 + \dots + w_n + \dots \quad (*)$$

with complex terms is said to be convergent if there exists the limit of the sequence of its partial sums.

The convergence of series (*) is equivalent to the convergence of the two series with real terms

$$u_1 + u_2 + \dots + u_n + \dots \quad \text{and} \quad v_1 + v_2 + \dots + v_n + \dots$$

where u_n and v_n are, respectively, the real and the imaginary parts of the n th term of the given complex series: $w_n = u_n + iv_n$.

If the series formed of the absolute values of the terms of series (*), that is the series

$$|w_1| + |w_2| + \dots + |w_n| + \dots$$

is convergent the original series is also convergent (in this case it is said to be *absolutely convergent*). This assertion is an immediate consequence of the inequalities $|u_n| \leq |w_n|$ and $|v_n| \leq |w_n|$.

Now, let $z = x + iy$ be a *complex variable*. Since we know (see Sec. 72) how a complex number is raised to the power n where n

is any positive integer we can consider the power function

$$w = z^n$$

(where n is a positive integer) of the complex argument z .

In contradistinction to complex functions of a real argument, in this case both *the independent variable and the function assume complex values*. In our course we shall not present the general theory of functions of a complex variable referring the reader to special books, for example to [3], [16], [19], and confining ourselves to the definitions of the simplest functions of a complex argument based on power series.

Consider a power series in the complex argument z :

$$a_0 + a_1 z + a_2 z^2 + \dots + a_n z^n + \dots \quad (**)$$

where the coefficients $a_0, a_1, \dots, a_n, \dots$ of the series are arbitrary complex numbers. It turns out that Abel's theorem (see Sec. 189) is readily generalized to series (**): *if series (*) is convergent at a point z_0 it is absolutely convergent for all z 's whose moduli are less than the modulus of z_0 : $|z| < |z_0|$* . From Abel's theorem it follows that there exists a real number R such that power series (**) is absolutely convergent for all the values of z satisfying the inequality $|z| < R$ and is divergent for all z 's such that $|z| > R$. The inequality $|z| < R$ specifies a circle of radius R with centre at the origin in the complex z -plane; the number R is called the *radius of convergence of series (**)* and the circle $|z| < R$ is its *circle of convergence*.

The radius of convergence may be equal to zero; then series (**) is only convergent at a single point, that is at the origin. A complex power series can also be convergent throughout the complex plane; then we say that its radius of convergence is equal to infinity.

The sum of a complex power series represents a function of a complex argument within its circle of convergence; such functions are termed *analytical*.

The power series

$$1 + z + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots \quad (***)$$

is convergent throughout the complex plane. This follows from the fact that it is convergent for any real value $z = N$ and hence is also convergent inside any circle with centre at the origin.

At the points of the real axis, i.e. for $z = x$, this series represents the real exponential function e^x :

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

Now we call, by definition, the sum of series (***) the (complex) exponential function e^z :

$$e^z = 1 + z + \frac{z^2}{2!} + \dots + \frac{z^n}{n!} + \dots$$

This function is defined throughout the complex plane.

The function e^z thus defined satisfies the basic condition

$$e^{z_1+z_2} = e^{z_1}e^{z_2}$$

which is fulfilled for the real exponential function for real $z_1 = x_1$ and $z_2 = x_2$. Indeed, on multiplying the series expressing e^{z_1} and e^{z_2} according to the rule stated in Sec. 186 we obtain

$$\begin{aligned} e^{z_1}e^{z_2} &= \left(1 + z_1 + \frac{z_1^2}{2!} + \frac{z_1^3}{3!} + \dots\right) \left(1 + z_2 + \frac{z_2^2}{2!} + \frac{z_2^3}{3!} + \dots\right) = \\ &= 1 + (z_1 + z_2) + \frac{1}{2!}(z_1^2 + 2z_1z_2 + z_2^2) + \\ &\quad + \frac{1}{3!}(z_1^3 + 3z_1^2z_2 + 3z_1z_2^2 + z_2^3) + \dots = \\ &= 1 + (z_1 + z_2) + \frac{(z_1 + z_2)^2}{2!} + \frac{(z_1 + z_2)^3}{3!} + \dots = e^{z_1+z_2} \end{aligned}$$

In particular, if $z_1 = x$ and $z_2 = iy$ we have $e^{x+iy} = e^x e^{iy}$. Using the series for e^z we can represent the second factor e^{iy} as the power series

$$e^{iy} = 1 + iy + \frac{(iy)^2}{2!} + \frac{(iy)^3}{3!} + \frac{(iy)^4}{4!} + \dots + \frac{(iy)^n}{n!} + \dots$$

Since $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etc. we have

$$e^{iy} = 1 + iy - \frac{y^2}{2!} - i \frac{y^3}{3!} + \frac{y^4}{4!} + i \frac{y^5}{5!} - \dots$$

Regrouping the terms (which is allowable since the series is absolutely convergent) we write

$$e^{iy} = \left(1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \dots\right) + i \left(y - \frac{y^3}{3!} + \frac{y^5}{5!} - \dots\right)$$

The expressions in the parentheses are, respectively, the power series representing $\cos y$ and $\sin y$, whence

$$e^{iy} = \cos y + i \sin y$$

We have thus arrived at *Euler's formula* which was derived in quite another way in Sec. 74 (see the footnote on p. 242). We remind the reader that Euler's formula makes it possible to compute the values of the function e^z for any exponent $z = x + iy$:

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y)$$

The trigonometric and hyperbolic functions of a complex variable are defined with the aid of the function e^z :

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin z = \frac{e^{iz} - e^{-iz}}{2i}$$

$$\cosh z = \frac{e^z + e^{-z}}{2}, \quad \sinh z = \frac{e^z - e^{-z}}{2}$$

All these functions are defined throughout the complex plane. If z takes on a real value x we obtain from these expressions the ordinary definitions of the hyperbolic functions of a real argument (see Sec. 23) and Euler's formulas expressing the cosine and the sine in terms of the exponential functions with imaginary exponents (Sec. 74).

Let the reader use these definitions to write down power series expansions of the complex trigonometric and hyperbolic functions and to check that they can be obtained from the known expansions in the real argument x by substituting z for x .

198*. Taylor's Series and Taylor's Formula for Functions of Two Independent Variables. In this section we shall briefly discuss power series expansions for functions of two variables; the main emphasis will be laid upon practical calculations, the questions of convergence being only limited to some general remarks.

Definition. A *power series in two variables* is an expression of the form

$$a_{00} + (a_{10}x + a_{01}y) + (a_{20}x^2 + a_{11}xy + a_{02}y^2) + \dots$$

$$\dots + (a_{n0}x^n + a_{n-1,1}x^{n-1}y + \dots + a_{0n}y^n) + \dots \quad (*)$$

The coefficients of this series are numbered with two indices: the coefficient in the product $x^k y^l$ where k and l are arbitrary integers is denoted a_{kl} . For the sake of convenience, the terms of the series are grouped so that each bracket involves the terms with constant value of the sum $k+l$, that is, is a homogeneous polynomial in x and y of degree $k+l$. We can also consider series in powers of the differences $x-x_0$ and $y-y_0$ for arbitrary x_0 and y_0 ; to this end we should replace x by $(x-x_0)$ and y by $(y-y_0)$ in series (*).

Series (*) is usually convergent in a domain symmetric about the coordinate axes (in particular, it may be convergent throughout the plane Oxy) and containing the point $(0, 0)$. *In every rectangle lying inside this domain series (*) is regularly convergent and its sum is a continuous function of the variables x and y . This function possesses partial derivatives of any order which can be found by means of term-by-term differentiation of the series.*

Now suppose that $f(x, y)$ is a function having the derivatives of all orders in a neighbourhood of the origin. Let us assume that it can be expanded into a power series of form (*). We

shall show that its coefficients can be found by means of the well-known method of undetermined coefficients.

Thus, let

$$f(x, y) = a_{00} + (a_{10}x + a_{01}y) + (a_{20}x^2 + a_{11}xy + a_{02}y^2) + \dots$$

Putting $x=0$ and $y=0$ we find $a_{00} = f(0, 0)$.

Let us compute the partial derivatives with respect to x and y :

$$\frac{\partial f}{\partial x} = a_{10} + 2a_{20}x + a_{11}y + \dots$$

$$\frac{\partial f}{\partial y} = a_{01} + a_{11}x + 2a_{02}y + \dots$$

The terms of these series designated by the dots contain variable factors of the form $x^k y^l$ with $k+l \geq 2$. Putting again $x=0$ and $y=0$ we obtain the expressions for a_{10} and a_{01} :

$$a_{10} = \left(\frac{\partial f}{\partial x} \right)_0, \quad a_{01} = \left(\frac{\partial f}{\partial y} \right)_0$$

Here and below the subscript "0" indicates that the derivatives are taken at the point $(0, 0)$.

Completely similarly we derive the formulas

$$a_{20} = \frac{1}{2!} \left(\frac{\partial^2 f}{\partial x^2} \right)_0, \quad a_{11} = \left(\frac{\partial^2 f}{\partial x \partial y} \right)_0, \quad a_{02} = \frac{1}{2!} \left(\frac{\partial^2 f}{\partial y^2} \right)_0, \text{ etc.}$$

Thus, the *Taylor series* of the function of two variables $f(x, y)$ has the form

$$\begin{aligned} f(x, y) = & f(0, 0) + \left(\frac{\partial f}{\partial x} \right)_0 x + \left(\frac{\partial f}{\partial y} \right)_0 y + \\ & + \frac{1}{2!} \left[\left(\frac{\partial^2 f}{\partial x^2} \right)_0 x^2 + 2 \left(\frac{\partial^2 f}{\partial x \partial y} \right)_0 xy + \left(\frac{\partial^2 f}{\partial y^2} \right)_0 y^2 \right] + \\ & + \frac{1}{3!} \left[\left(\frac{\partial^3 f}{\partial x^3} \right)_0 x^3 + 3 \left(\frac{\partial^3 f}{\partial x^2 \partial y} \right)_0 x^2 y + 3 \left(\frac{\partial^3 f}{\partial x \partial y^2} \right)_0 xy^2 + \left(\frac{\partial^3 f}{\partial y^3} \right)_0 y^3 \right] + \dots \end{aligned}$$

The structure of the coefficients in the subsequent terms is quite clear.

Taking the difference between the function $f(x, y)$ and the partial sum of its Taylor series including all the terms up to the degree n inclusive we obtain the *remainder term of Taylor's series*. Like in the case of a function of one variable, it can be proved that the remainder is obtained if we take the group of terms of degree $n+1$ and replace the values of the corresponding derivatives taken at the point $(0, 0)$ by their values at an intermediate point (ξ, η) lying on the line segment joining the point (x, y) to the origin.

Replacing the remainder term by its expression described above we obtain *Taylor's formula with the remainder for the function of two variables*. Let the reader write down this formula for $n=2$.

It is clear that Taylor's series converges to the function $f(x, y)$ only if the remainder tends to zero as $n \rightarrow \infty$.

If a function $f(x, y)$ is expanded into Taylor's series in the vicinity of a point (x_0, y_0) the values of the partial derivatives are taken at that point and x and y are replaced, respectively, by $(x-x_0)$ and $(y-y_0)$.

In practice, when expanding a function of two variables into Taylor's series we can use the known expansions of functions of one variable.

Examples. (1) Let us expand into Taylor's series the function $e^x \sin y$ in the neighbourhood of the point $(0, 0)$. On replacing e^x and $\sin y$ by their expansions and multiplying the series we obtain

$$\begin{aligned} e^x \sin y &= \left(1 + x + \frac{x^2}{2!} + \dots\right) \left(y - \frac{y^3}{3!} + \frac{y^5}{5!} - \dots\right) = \\ &= y + xy + \left(\frac{x^2 y}{2!} - \frac{y^3}{3!}\right) + \dots \end{aligned}$$

The same result is of course obtained if we use the general formula.

(2) Let us expand into a power series the function $z = x^y$ in the vicinity of the point $(1, 1)$. The computation of the partial derivatives yields

$$\begin{aligned} z'_x &= yx^{y-1}, & z'_y &= x^y \ln x \\ z''_{xx} &= y(y-1)x^{y-2}, & z''_{xy} &= x^{y-1} + yx^{y-1} \ln x, \\ z''_{yy} &= x^y \ln^2 x \end{aligned}$$

.

On computing the values of the derivatives at the point $(1, 1)$ and substituting them into the general formula we receive

$$x^y = 1 + (x-1) + (x-1)(y-1) + \dots$$

For instance, it follows that $(1.04)^{1.03} \approx 1 + 0.04 + 0.04 \cdot 0.03 = 1.0412$.

QUESTIONS

1. What do we call a numerical series? What is the general term of the series?

2. State the definition of the sum of a numerical series. State the definitions of a convergent and of a divergent series and give illustrative examples.

3. What is the necessary condition for the convergence of a series? Give an example showing that this condition is not sufficient.

4. State the simplest test for the divergence of a series.

5. State and prove the comparison tests for two positive series.

6. State and prove D'Alembert's test.

7. State and prove Cauchy's integral test. Prove that the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent for $p > 1$.
8. What is an alternating series? State and prove Leibniz' test for such a series.
9. State the sufficient test for the convergence of a series with arbitrary terms. Prove this test.
10. What is an absolutely convergent series? What do we call a conditionally convergent series? Give examples of absolute and conditional convergence.
11. What do we call a functional series? State the definition of the domain of convergence of a functional series.
12. State the definition of a regularly convergent functional series.
13. Enumerate the properties of regularly convergent functional series.
14. What do we call a power series?
15. State and prove Abel's theorem. What is the interval of convergence of a power series? What is the radius of convergence?
16. Give examples of power series whose radii of convergence are equal (1) to zero, (2) to infinity and (3) to a finite nonzero number.
17. Prove the lemmas on power series and enumerate the properties of power series implied by them.
18. State the problem of expanding a function $f(x)$ into a power series.
19. What do we call the Taylor series of a function $f(x)$? How are Taylor's coefficients determined?
20. What is the remainder term of Taylor's series?
21. State and prove Lagrange's theorem on the form of the remainder of Taylor's series.
22. What is Taylor's polynomial? Write down Taylor's formula with the remainder after the n th term.
23. What do we call Maclaurin's series?
24. Derive the expansions into Maclaurin's series of the functions e^x , $\sin x$, $\cos x$, $(1+x)^m$, $\ln(1+x)$ and $\arctan x$. For what intervals are these expansions valid?
25. Enumerate the techniques used for estimating the remainder in the approximate calculation of the value of a function.
26. How are power series applied to integrating functions? Give examples.
27. How do we use power series for solving differential equations? Give examples.
- 28*. State the definition of the limit of a sequence of complex numbers. What is a convergent series with complex terms?

29*. State the definition of the complex exponential function e^z by means of a power series and prove the basic property of this function.

30*. Apply power series to derive Euler's formula.

31*. State the definitions of the trigonometric functions $\cos z$, $\sin z$ and the hyperbolic functions $\cosh z$, $\sinh z$ of the complex argument z .

32*. What is a power series in two variables x and y ?

33*. Derive the formula for the expansion of a function of two independent variables into Taylor's series with the aid of the method of undetermined coefficients.

hold where r_1 and r_2 are *positive integers*. (The number T is the *primitive period*, that is the *smallest* one, if the numbers r_1 and r_2 are coprime.) It follows that in this case the ratio of the frequencies is equal to the ratio of these integers:

$$\frac{\omega_2}{\omega_1} = \frac{r_2}{r_1}$$

Consequently, the frequencies ω_1 and ω_2 are *commensurable*. If the frequencies are *incommensurable* the resultant vibration is not periodic. If the frequencies are commensurable we can put

$$\omega_1 = r_1 \omega, \quad \omega_2 = r_2 \omega$$

and write the superposition of the vibrations as

$$s = A_1 \sin(r_1 \omega t + \varphi_1) + A_2 \sin(r_2 \omega t + \varphi_2)$$

This is a periodic function with period

$$T = \frac{2\pi}{\omega}$$

Indeed, the function s does not change its value if the increment $\frac{2\pi}{\omega}$ is added to any value of t ; if in addition, the integers r_1 and r_2 are coprime the number $\frac{2\pi}{\omega}$ is the *primitive period*.

For simplicity, let us assume that $\omega = 1$, i.e. $T = 2\pi$. (The case when ω is an arbitrary number is reduced to the former if we change the scale along the t -axis, that is put $\omega t = t'$.)

The graph of a periodic function of the form

$$A_1 \sin(r_1 t + \varphi_1) + A_2 \sin(r_2 t + \varphi_2)$$

may differ considerably in its shape from the graphs of simple harmonics (e.g. see Fig. 29 on p. 64). Generally, a superposition of several simple harmonics with commensurable frequencies, i.e. an expression of the form

$$y = A_1 \sin(r_1 t + \varphi_1) + A_2 \sin(r_2 t + \varphi_2) + \dots + A_n \sin(r_n t + \varphi_n)$$

provides a wide variety of periodic functions for different values of the parameters A_k , φ_k and the integers r_k and n . From the point of view of mechanics this means that the superpositions of simple harmonic motions form various periodic motions essentially differing from simple harmonic motions.

Therefore it appears natural to pose the reverse problem, that is, the question as to *whether it is possible to choose the constituent simple harmonic motions in such a way that their superposition is a given periodic vibration or, in other words, whether it is possible to represent any periodic motion as a superposition of simple harmonic vibrations*.

It turns out that, generally speaking, this cannot be achieved if we use only finite sums of simple harmonics, but if infinite sums, that is series, of simple harmonics are allowed then practically every periodic function can be expanded into simple harmonics. This is the problem we shall be concerned with in this chapter.

II. Trigonometric series.

Definition. A functional series of the form

$$\frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \dots \\ \dots + a_n \cos nx + b_n \sin nx + \dots$$

whose terms are constant factors multiplied by sines and cosines of integer multiples of the values of the argument x is called a *trigonometric series*.

The constant factors a_i and b_i ($i=1, 2, \dots$) are termed the *coefficients* of the trigonometric series. The constant term is written as $\frac{a_0}{2}$ to unify the formulas derived later.

In the abbreviated notation we shall write a trigonometric series in the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad (*)$$

Trigonometric series (*) can be represented as a sum of simple harmonics only involving the sines (or the cosines). To this end we combine the summands $\cos nx$ and $\sin nx$ ($n=1, 2, \dots$) with the same frequencies and transform the combinations

$$a_n \cos nx + b_n \sin nx$$

by putting $a_n = A_n \sin \varphi_n$ and $b_n = A_n \cos \varphi_n$. This results in

$$a_n \cos nx + b_n \sin nx = A_n \sin (nx + \varphi_n)$$

This form of representation is convenient when it is necessary to determine the amplitude and the initial phase of the n th harmonics. In this equality

$$A_n = \sqrt{a_n^2 + b_n^2} \text{ and } \tan \varphi_n = \frac{a_n}{b_n}$$

Series (*) then takes the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} A_n \sin (nx + \varphi_n)$$

The n th harmonic can also be written as

$$a_n \cos nx + b_n \sin nx = A_n \cos (nx - \varphi_n)$$

where the amplitude A_n is again equal to $\sqrt{a_n^2 + b_n^2}$ but the initial phase has another expression: $\tan \varphi_n = \frac{b_n}{a_n}$.

All the terms of trigonometric series (*) are periodic with period $T = 2\pi$. Therefore, if this series is convergent in the interval $[-\pi, \pi]$ (of course, the substitution of the values $x = -\pi$ and $x = \pi$ into the trigonometric series results in the same numerical series), it is also convergent throughout the Ox -axis and its sum is a periodic function with period 2π . We have chosen the interval $[-\pi, \pi]$ for definiteness; it can also be replaced by any other interval of length 2π , e.g. by the interval $[0, 2\pi]$.

Let us derive some auxiliary relations¹ which will be applied to determining the coefficients a_i and b_i (the number k and p in the equalities below are arbitrary nonnegative integers):

$$(1) \int_{-\pi}^{\pi} \sin kx \, dx = 0 \quad \text{for any } k.$$

$$(2) \int_{-\pi}^{\pi} \cos kx \, dx = 0 \quad \text{if } k \neq 0.$$

Indeed, we have $\int_{-\pi}^{\pi} \sin kx \, dx = -\frac{1}{k} \cos kx \Big|_{-\pi}^{\pi} = 0$ since $\cos x$ is an even function ($\cos k\pi = \cos(-k\pi)$); we also have $\int_{-\pi}^{\pi} \cos kx \, dx = \frac{1}{k} \sin kx \Big|_{-\pi}^{\pi} = 0$ since $\sin k\pi = \sin(-k\pi) = 0$.

$$(3) \int_{-\pi}^{\pi} \cos kx \sin px \, dx = 0 \quad \text{for any } k \text{ and } p.$$

$$(4) \int_{-\pi}^{\pi} \cos kx \cos px \, dx = \begin{cases} 0 & \text{if } k \neq p; \\ \pi & \text{if } k = p \neq 0. \end{cases}$$

$$(5) \int_{-\pi}^{\pi} \sin kx \sin px \, dx = \begin{cases} 0 & \text{if } k \neq p, \\ \pi & \text{if } k = p \neq 0. \end{cases}$$

These relations are readily proved with the help of the well-known trigonometric formulas

$$\cos kx \sin px = \frac{1}{2} [\sin(k+p)x + \sin(p-k)x]$$

$$\cos kx \cos px = \frac{1}{2} [\cos(k+p)x + \cos(k-p)x]$$

¹ In Sec. 205 the role of these relations will be elucidated from a more general point of view.

and

$$\sin kx \sin px = \frac{1}{2} [\cos (k-p)x - \cos (k+p)x]$$

If $k \neq p$ the substitution of the right-hand expressions for the products of trigonometric functions into integrals (3), (4) and (5) splits each of the integrals into two integrals. The latter integrals are equal to zero by virtue of formulas (1) and (2) because $k+p$ and $k-p$ are nonzero integers.

If $k=p$ we have

$$\int_{-\pi}^{\pi} \cos^2 kx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} (1 + \cos 2kx) \, dx = \frac{1}{2} x \Big|_{-\pi}^{\pi} = \pi$$

(the integral $\int_{-\pi}^{\pi} \cos 2kx \, dx$ is equal to zero according to the same formula (2)).

Similarly,

$$\int_{-\pi}^{\pi} \sin^2 kx \, dx = \frac{1}{2} \int_{-\pi}^{\pi} (1 - \cos 2kx) \, dx = \pi$$

Thus, relations (1)-(5) have been completely proved.

200. Fourier Series. Let $f(x)$ be a function defined in the closed interval $[-\pi, \pi]$. Let us suppose that this function can be expanded into a convergent trigonometric series, i.e. can be represented in the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad (*)$$

We shall also suppose that series (*) can be integrated term-wise, that is the integral of the sum of the series $f(x)$ is equal to the sum of the integrals of the terms of the series.

On integrating both sides of expansion (*) from $-\pi$ to π we obtain

$$\int_{-\pi}^{\pi} f(x) \, dx = \int_{-\pi}^{\pi} \frac{a_0}{2} \, dx = \pi a_0$$

By virtue of formulas (1) and (2) of Sec. 199 the integrals of all the other terms of the series are equal to zero. It follows that

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \, dx \quad (1)$$

To determine the coefficient a_k where k is an arbitrary positive integer we multiply both members of equality (*) by $\cos kx$ and integrate the resultant relation from $-\pi$ to π (here and henceforth we also suppose that the term-by-term integration is permissible):

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = \frac{a_0}{2} \int_{-\pi}^{\pi} \cos kx dx + \\ + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} \cos nx \cos kx dx + b_n \int_{-\pi}^{\pi} \sin nx \cos kx dx \right)$$

All the integrals on the right-hand side except the one with coefficient a_k vanish in accordance with formulas (2), (3) and (4) of Sec. 199. The integral with the coefficient a_k is equal to π whence we find this coefficient:

$$\int_{-\pi}^{\pi} f(x) \cos kx dx = a_k \int_{-\pi}^{\pi} \cos^2 kx dx = \pi a_k$$

and

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx \quad (2)$$

Formula (1) is obtained from formula (2) for $k=0$.

In order to find the coefficient b_k we multiply both members of equality (*) by $\sin kx$ and integrate the result from $-\pi$ to π :

$$\int_{-\pi}^{\pi} f(x) \sin kx dx = \frac{a_0}{2} \int_{-\pi}^{\pi} \sin kx dx + \\ + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} \cos nx \sin kx dx + b_n \int_{-\pi}^{\pi} \sin nx \sin kx dx \right)$$

By formulas (1), (3) and (5) of Sec. 199, all the integrals on the right except the one with coefficients b_k are equal to zero. The

latter integral is equal to $\int_{-\pi}^{\pi} \sin kx \sin kx dx = \pi$. Consequently,

$$\int_{-\pi}^{\pi} f(x) \sin kx dx = \pi b_k \quad \text{and} \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx \quad (3)$$

Thus, proceeding from the assumption that the function $f(x)$ can be expanded into a trigonometric series (*) we managed to determine all its coefficients.

Now, let $f(x)$ be an arbitrary function defined in the interval $[-\pi, \pi]$ for which we only suppose that all the integrals below exist (this function may have points of discontinuity).

Definition. The numbers a_n and b_n specified by the formulas

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx \quad \text{and} \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx$$

are called the *Fourier coefficients* of the function $f(x)$. The series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is called the *Fourier series* of the function $f(x)$.

We cannot of course assert that the Fourier series of an arbitrary function $f(x)$ is convergent and that its sum (in case it is convergent) is equal to the function $f(x)$. (A similar situation arises in the theory of Taylor's series; see Sec. 191.)

Before proceeding to the key conditions guaranteeing the convergence of the Fourier series of a given function $f(x)$ and the coincidence of $f(x)$ with the sum of the series (in case it is convergent) we make some additional remarks.

As was mentioned in Sec. 199, II, if a trigonometric series is convergent its sum is a periodic function with period 2π . Therefore, when speaking about the expansion of a function $f(x)$ into

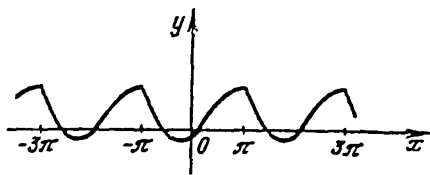


Fig. 251

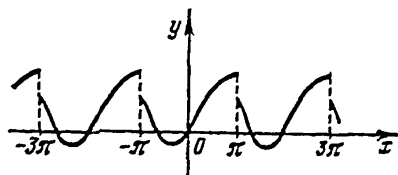


Fig. 252

Fourier's series, we shall suppose that this function is originally defined in the interval $[-\pi, \pi]$ and then periodically extended with period 2π to the whole x -axis. This is of course only possible if $f(\pi) = f(-\pi)$ (see Fig. 251). If otherwise, i.e. if $f(\pi) \neq f(-\pi)$, we can leave the values of the function $f(x)$ unchanged in the interval $[-\pi, \pi)$ and extend it periodically with period 2π to the entire Ox -axis. This results in the appearance of discontinuities at the points $\pi + 2k\pi$ where $k=0, \pm 1, \pm 2, \dots$ (Fig. 252).

Let us state the definition of a *piecewise smooth* function:

A function $f(x)$ is said to be *smooth* in a closed interval $[a, b]$ if it is continuous in this interval together with its first derivative $f'(x)$.

A function $f(x)$ is said to be *piecewise smooth* in an interval $[a, b]$ if the latter can be divided into a finite number of closed subintervals in each of which the function $f(x)$ is smooth.

The graph of a smooth function is a *smooth curve* (see Sec. 66); there are no corner points and cusps on such a curve. The graph of a piecewise smooth function consists of a finite number of smooth arcs; such a line is spoken of as a *piecewise smooth curve*.

These definitions imply that a piecewise smooth function defined in an interval $[a, b]$ can have at most a finite number of points of discontinuity of the first kind. For instance, the function whose graph is shown in Fig. 253 is piecewise smooth in the interval $[a, b]$.

We remind the reader that at a point of discontinuity of the first kind the function possesses right-hand and left-hand limits which we conditionally denote $f(x-0)$ and $f(x+0)$.

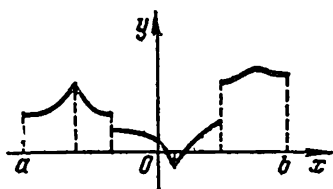


Fig. 253

Now we state without proof the key theorem on the possibility of expanding a given function $f(x)$ into Fourier's series:

Theorem. If the function $f(x)$ is piecewise smooth in the interval $[-\pi, \pi]$ its Fourier series converges to the function $f(x)$ at all the interior points of the interval at which the function is continuous.

If x_0 is a point of discontinuity of the function $f(x)$ ($-\pi < x_0 < \pi$) the Fourier series of the function converges at that point x_0 to the arithmetic mean of its right-hand and left-hand limits, that is to $\frac{f(x_0-0) + f(x_0+0)}{2}$.

At both end points of the interval $[-\pi, \pi]$ the sum of the series is equal to the arithmetic mean of the right-hand limit at the point $x = -\pi$ and the left-hand limit at the point $x = \pi$, that is to the number

$$\frac{f(-\pi+0) + f(\pi-0)}{2}$$

In particular, if $f(x)$ is a smooth function and its values at the end points of the interval $[-\pi, \pi]$ coincide its Fourier series is convergent to $f(x)$ throughout the closed interval $[-\pi, \pi]$.

The assertions of the key theorem remain true under the weaker condition that the function $f(x)$ has only a finite number of maxima and minima in the interval $[-\pi, \pi]$ and is continuous in that interval except a finite number of points of discontinuity of the first kind (this is known as *Dirichlet's condition*).

In mathematical analysis and its applications we usually deal with piecewise smooth functions or functions satisfying Dirichlet's condition which are therefore representable by means of their Fourier series.

It should be noted that the requirements imposed on a function when it is expanded into Fourier's series are essentially

weaker than when it is expanded into a power series. Indeed, a function representable by Taylor's series is not only continuous throughout the interval of convergence but also infinitely differentiable while it only suffices that the function should be piecewise smooth when it is expanded into Fourier's series.

For a function which can be expanded into Fourier's series we can obtain an accurate approximation by taking a sufficient number of terms of its Fourier series, that is by replacing the function by a partial sum of that series. By analogy with Taylor's series, we call such a finite sum a *Fourier polynomial* of the given function:

$$f(x) \approx \frac{a_0}{2} + \sum_{n=1}^k (a_n \cos nx + b_n \sin nx)$$

In connection with such an approximation it is natural to pose the problem of estimating the error of this approximate inequality. But this is a more complicated question compared with the analogous problem in the theory of Taylor's series and we shall not dwell on it.

The theory of Fourier's series appeared in connection with concrete problems of mechanics and physics. Fourier's series have many important applications in various divisions of mathematics and provide particularly convenient methods for the solution of many problems of mathematical physics¹.

201. Expanding Even and Odd Functions into Fourier Series. Fourier Series for an Arbitrary Interval.

1. Expanding even and odd functions. Suppose that a function $f(x)$ expanded into Fourier's series is *even*. Then the functions $f(x) \sin kx$, $k=1, 2, \dots$, are odd, and all the coefficients b_k are equal to zero as integrals of odd functions over the interval $[-\pi, \pi]$ symmetric about the origin (see Sec. 94). Consequently, the Fourier series of an even function only involves cosines; it has the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx \, dx$$

since $f(x) \cos nx$ is an even function.

Now let $f(x)$ be an *odd* function. Then the functions $f(x) \cos kx$ are odd, and all the coefficients a_k are equal to zero. Consequ-

¹ In particular, trigonometric series play an important role in the study of nonsinusoidal periodic electric currents in electrical engineering.

ently, the Fourier series of an odd function involves only sines and is written as

$$\sum_{n=1}^{\infty} b_n \sin nx$$

where

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \, dx = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx$$

since $f(x) \sin nx$ is an even function.

II. Expanding functions with arbitrary period. Let us consider the problem of expanding into Fourier's series a function $f(x)$ defined in an interval $[-l, l]$ where l is an arbitrary positive number.

If, in the interval $[-l, l]$, the function $f(x)$ satisfies the conditions of the theorem stated in Sec. 200 its expansion can be readily obtained by changing variable according to the formula $x' = \frac{\pi}{l}x$. Then we pass to the function $f\left(\frac{l}{\pi}x'\right)$, and when the variable x runs through the interval $[-l, l]$ the new variable x' runs through the interval $[-\pi, \pi]$. The expansion of the new function into Fourier's series has the form

$$f\left(\frac{l}{\pi}x'\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx' + b_n \sin nx')$$

Now, returning to the old variable, we arrive at the expansion of the given function $f(x)$:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l} \right)$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{l}{\pi}x'\right) \cos nx' \, dx' = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi x}{l} \, dx$$

and

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f\left(\frac{l}{\pi}x'\right) \sin nx' \, dx' = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi x}{l} \, dx$$

(in the integrals on the left we have made the substitution $\frac{l}{\pi}x' = x$).

The sum of the Fourier series thus constructed is a periodic function of x with period $T = 2l$.

weaker than when it is expanded into a power series. Indeed, a function representable by Taylor's series is not only continuous throughout the interval of convergence but also infinitely differentiable while it only suffices that the function should be piecewise smooth when it is expanded into Fourier's series.

For a function which can be expanded into Fourier's series we can obtain an accurate approximation by taking a sufficient number of terms of its Fourier series, that is by replacing the function by a partial sum of that series. By analogy with Taylor's series, we call such a finite sum a *Fourier polynomial* of the given function:

$$f(x) \approx \frac{a_0}{2} + \sum_{n=1}^k (a_n \cos nx + b_n \sin nx)$$

In connection with such an approximation it is natural to pose the problem of estimating the error of this approximate inequality. But this is a more complicated question compared with the analogous problem in the theory of Taylor's series and we shall not dwell on it.

The theory of Fourier's series appeared in connection with concrete problems of mechanics and physics. Fourier's series have many important applications in various divisions of mathematics and provide particularly convenient methods for the solution of many problems of mathematical physics¹.

201. Expanding Even and Odd Functions into Fourier Series. Fourier Series for an Arbitrary Interval.

1. Expanding even and odd functions. Suppose that a function $f(x)$ expanded into Fourier's series is *even*. Then the functions $f(x) \sin kx$, $k=1, 2, \dots$, are odd, and all the coefficients b_k are equal to zero as integrals of odd functions over the interval $[-\pi, \pi]$ symmetric about the origin (see Sec. 94). Consequently, the Fourier series of an even function only involves cosines; it has the form

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx$$

where

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \, dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx \, dx$$

since $f(x) \cos nx$ is an even function.

Now let $f(x)$ be an *odd* function. Then the functions $f(x) \cos kx$ are odd, and all the coefficients a_k are equal to zero. Consequ-

¹ In particular, trigonometric series play an important role in the study of nonsinusoidal periodic electric currents in electrical engineering.

Examples. (1) Let us expand into Fourier's series the function $f(x)$ defined in the interval $[-\pi, \pi]$ by the conditions

$$f(x) = \begin{cases} -1 & \text{for } -\pi \leq x < 0 \\ 1 & \text{for } 0 < x \leq \pi \end{cases}$$

This function is discontinuous at the point $x=0$. In Fig. 254 the graph of this function is shown as extended periodically with period 2π to the whole Ox -axis.

The function clearly satisfies the conditions of the key theorem and, since it is odd, its all coefficients a_n are zero.

Let us determine the coefficients b_n . By Sec. 201, we have

$$b_n = \frac{2}{\pi} \int_0^{\pi} 1 \cdot \sin nx \, dx =$$

$$= -\frac{2}{\pi n} \cos nx \Big|_0^{\pi} = -\frac{2}{\pi n} (\cos n\pi - 1)$$

that is

$$b_n = \frac{2}{\pi n} (1 - \cos n\pi).$$

Since $\cos n\pi = (-1)^n$, we receive, by putting, in succession, $n=1, 2, 3, \dots$, the numerical values of the coefficients b_k :

$$b_1 = \frac{4}{\pi}, \quad b_2 = 0, \quad b_3 = \frac{4}{3\pi}, \quad b_4 = 0, \quad b_5 = \frac{4}{5\pi}, \text{ etc.}$$

Thus,

$$f(x) = \frac{4}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \dots \right)$$

According to the key theorem, the sum of the series on the right is equal to 1 for $0 < x < \pi$ and to -1 for $-\pi < x < 0$. At the point $x=0$ and also at the points $\pm\pi$ the value of the sum is equal to zero, which is again coherent with the key theorem. The behaviour of the series is quite the same in every interval obtained by shifting the original interval $[-\pi, \pi]$ to the left or to the right by an integral number of periods.

It is interesting to compare the graphs of trigonometric polynomials, equal to partial sums of this series, which approximate the given function, with the graph of the function $f(x)$. In Fig. 255 we see the graph of the function $f(x)$ and the graphs of the first and second partial sums shown in continuous line and the graphs of the constituent harmonics given in dotted line.

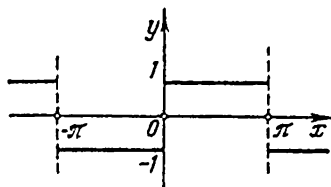


Fig. 254

The frequencies of the harmonics the series is formed of are $\omega_n = \frac{\pi n}{l}$, $n = 1, 2, \dots$, the lowest frequency being equal to $\omega_1 = \frac{\pi}{l}$.

III. Expanding functions into half-range Fourier series. Consider a function $f(x)$ defined in an interval $[0, l]$ and suppose that it is required to expand it into a trigonometric series. Since there always exists a change of variable reducing the case of the interval $[0, l]$ to the standard interval $[0, \pi]$ we shall limit ourselves to the latter. The given function $f(x)$ (defined, as we agreed, in the interval $[0, \pi]$), can be *extended in an arbitrary fashion* to the interval $[-\pi, 0)$ so that the new function $F(x)$ defined in the interval and coinciding with $f(x)$ in the interval $[0, \pi]$ satisfies all the requirements of the key theorem in the whole interval $[-\pi, \pi]$. On expanding the function $F(x)$ into Fourier's series in the interval $[-\pi, \pi]$ we obtain the sought-for trigonometric series representing the original function $f(x)$ in the interval $[0, \pi]$. Since the extension of $f(x)$ to $F(x)$ can be performed in an arbitrary manner there exist infinitely many such trigonometric series.

In particular, the function $f(x)$ can be extended to the interval $[-\pi, 0)$ as an *even function* defined in the whole interval $[-\pi, \pi]$; the graph of the new function $F(x)$ is then obtained if we extend the graph of $f(x)$ symmetrically with respect to the Oy -axis. The resulting function $F(x)$ is even, and its Fourier series only involves cosines. Similarly, the function $f(x)$ can be extended to the interval $[-\pi, 0)$ as an *odd function* defined in the interval $[-\pi, \pi]$; then the graph of the function $F(x)$ is the extension of the graph of $f(x)$ symmetric with respect to the origin, and $F(x)$ is an odd function whose Fourier series only involves sines. Such series containing either $\sin kx$ or $\cos kx$, $k = 1, 2, \dots$, are termed, respectively, *Fourier's sine series* or *Fourier's cosine series*. We also say in these cases that the given function $f(x)$ is expanded into a *half-range* (or *incomplete*) *Fourier series*.

Note that if the function $f(x)$ is continuous at the point $x = 0$ its even extension $F(x)$ is also continuous at this point. If $f(x)$ is extended as an odd function the resulting function $F(x)$ is continuous at the origin only if $f(0) = 0$; if $f(0) \neq 0$ the function $F(x)$ has a discontinuity of the first kind at the point $x = 0$.

As was said, if a function $f(x)$ defined in the interval $[0, \pi]$ can be expanded into a trigonometric series there exist an infinite number of such expansions. *In applications the most important role is played by the expansions into Fourier's sine and cosine series.*

Examples. (1) Let us expand into Fourier's series the function $f(x)$ defined in the interval $[-\pi, \pi]$ by the conditions

$$f(x) = \begin{cases} -1 & \text{for } -\pi \leq x < 0 \\ 1 & \text{for } 0 < x \leq \pi \end{cases}$$

This function is discontinuous at the point $x=0$. In Fig. 254 the graph of this function is shown as extended periodically with period 2π to the whole Ox -axis.

The function clearly satisfies the conditions of the key theorem and, since it is odd, its all coefficients a_n are zero.

Let us determine the coefficients b_n . By Sec. 201, we have

$$b_n = \frac{2}{\pi} \int_0^{\pi} 1 \cdot \sin nx \, dx =$$

$$= -\frac{2}{\pi n} \cos nx \Big|_0^{\pi} = -\frac{2}{\pi n} (\cos n\pi - 1)$$

that is

$$b_n = \frac{2}{\pi n} (1 - \cos n\pi).$$

Since $\cos n\pi = (-1)^n$, we receive, by putting, in succession, $n=1, 2, 3, \dots$, the numerical values of the coefficients b_k :

$$b_1 = \frac{4}{\pi}, \quad b_2 = 0, \quad b_3 = \frac{4}{3\pi}, \quad b_4 = 0, \quad b_5 = \frac{4}{5\pi}, \text{ etc.}$$

Thus,

$$f(x) = \frac{4}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \dots \right)$$

According to the key theorem, the sum of the series on the right is equal to 1 for $0 < x < \pi$ and to -1 for $-\pi < x < 0$. At the point $x=0$ and also at the points $\pm\pi$ the value of the sum is equal to zero, which is again coherent with the key theorem. The behaviour of the series is quite the same in every interval obtained by shifting the original interval $[-\pi, \pi]$ to the left or to the right by an integral number of periods.

It is interesting to compare the graphs of trigonometric polynomials, equal to partial sums of this series, which approximate the given function, with the graph of the function $f(x)$. In Fig. 255 we see the graph of the function $f(x)$ and the graphs of the first and second partial sums shown in continuous. The graphs of the constituent harmonics given in dotted.

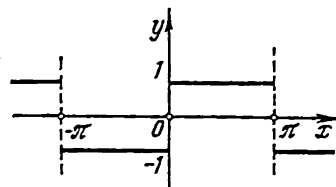


Fig. 254

Proceeding from the expansion of the function $f(x)$ found above we can readily write down the expansion of the function

$$\varphi(x) = \begin{cases} a & \text{for } -\pi \leq x < 0 \\ b & \text{for } 0 < x \leq \pi \end{cases}$$

where a and b are arbitrary numbers. To this end we use the obvious relationship between the old function $f(x)$ and the new

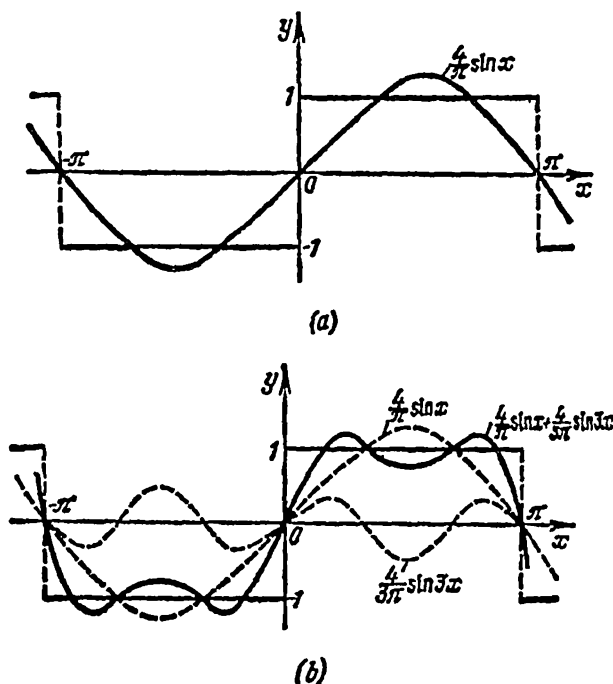


Fig. 255

function $\varphi(x)$:

$$f(x) = \frac{\varphi(x) - \frac{a+b}{2}}{\frac{b-a}{2}}$$

(This relationship means that the graph of $\varphi(x)$ can be obtained from that of $f(x)$ if we translate the origin to the point $(0, \frac{a+b}{2})$ and change the scale along Oy .) It follows that

$$\varphi(x) = \frac{a+b}{2} + \frac{2(b-a)}{\pi} \left(\sin x + \frac{\sin 3x}{3} + \dots \right)$$

(2) Let us find the expansion of the function

$$f(x) = |x|, \quad -\pi \leq x \leq \pi$$

into Fourier's series:

The periodic extension with period 2π of the graph of this function is shown in Fig. 256. The function $f(x)$ being *even*, all the coefficients b_n are equal to zero. Using the formulas of Sec. 201, I we get

$$a_0 = \frac{2}{\pi} \int_0^{\pi} x dx = \frac{2}{\pi} \frac{x^2}{2} \Big|_0^{\pi} = \pi, \quad a_n = \frac{2}{\pi} \int_0^{\pi} x \cos nx dx$$

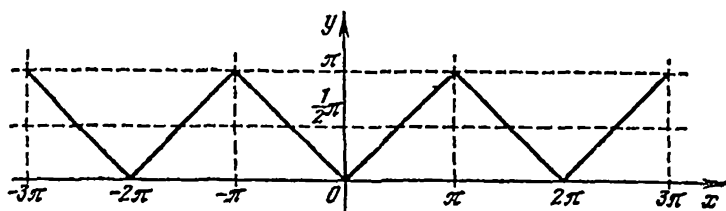


Fig. 256

The integration by parts leads to

$$a_n = \frac{2}{\pi} \left[\frac{x \sin nx}{n} \Big|_0^{\pi} - \int_0^{\pi} \frac{\sin nx}{n} dx \right] = \frac{2}{\pi n^2} \cos nx \Big|_0^{\pi} = \frac{2}{\pi n^2} (\cos \pi n - 1)$$

whence

$$a_1 = -\frac{4}{\pi}, \quad a_2 = 0, \quad a_3 = -\frac{4}{3^2\pi}, \quad a_4 = 0, \quad a_5 = -\frac{4}{5^2\pi}, \text{ etc.}$$

Thus,

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi} \left(\cos x + \frac{1}{3^2} \cos 3x + \frac{1}{5^2} \cos 5x + \dots \right)$$

The approximations to the given function with the partial sums $\frac{1}{2}\pi - \frac{4}{\pi} \cos x$ and $\frac{1}{2}\pi - \frac{4}{\pi} \cos x - \frac{4}{\pi} \frac{\cos^3 x}{3^2}$ are shown in Fig. 257.

Putting $x=0$ in the series obtained we receive an interesting numerical series representing the number π :

$$0 = \frac{\pi}{2} - \frac{4}{\pi} \left(1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots \right)$$

whence

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots$$

This formula implies some other useful relations. Let us put

$$\sigma = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \dots, \quad \sigma_1 = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots = \frac{\pi^2}{8}$$

and

$$\sigma_2 = \frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \dots = \frac{1}{4} \left(1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots \right)$$

We have $\sigma = \sigma_1 + \sigma_2$ and $\sigma_2 = \frac{1}{4} \sigma = \frac{1}{4} (\sigma_1 + \sigma_2)$, that is

$$3\sigma_2 = \sigma_1, \text{ which yields } \sigma_2 = \frac{\pi^2}{24}$$

Thus, we obtain

$$\sigma = \frac{\pi^2}{8} + \frac{\pi^2}{24} = \frac{\pi^2}{6}, \text{ i.e. } \frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \frac{1}{5^2} + \dots$$

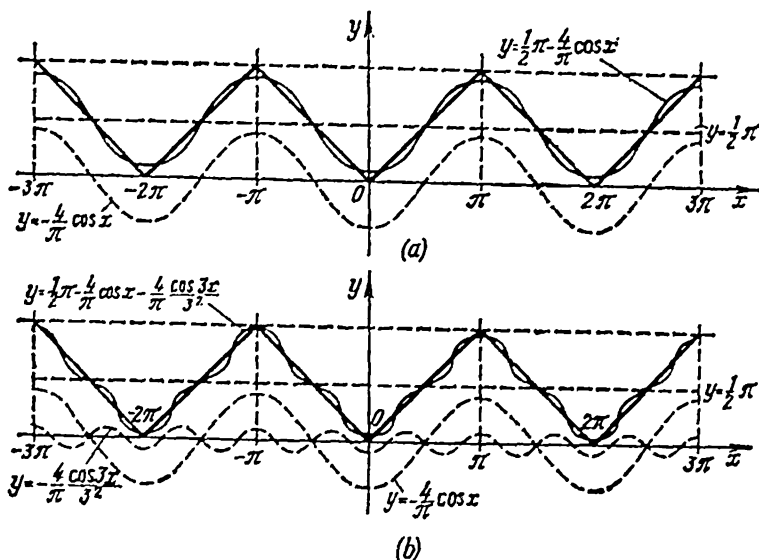


Fig. 257

In Sec. 184 we established the convergence of the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ but its sum was written without proof. Now we have found this sum.

(3) Let us expand into Fourier's series the function

$$f(x) = x, \quad -\pi \leq x \leq \pi$$

The graph of this function, when extended periodically to the whole x -axis, consists of parallel line segments (see Fig. 258). The function being odd, it suffices to determine b_n . We have

$$\begin{aligned} b_n &= \frac{2}{\pi} \int_0^{\pi} x \sin nx \, dx = \frac{2}{\pi} \left[-\frac{x \cos nx}{n} \Big|_0^{\pi} + \int_0^{\pi} \frac{\cos nx}{n} \, dx \right] = \\ &= \frac{2}{\pi} \left[\frac{-\pi \cos n\pi}{n} + \frac{\sin nx}{n^2} \Big|_0^{\pi} \right] = -\frac{2 \cos n\pi}{n} \end{aligned}$$

Putting in succession $n = 1, 2, 3, \dots$ we obtain $b_1 = 2$, $b_2 = -\frac{2}{2}$, $b_3 = \frac{2}{3}$, $\dots$

Thus,

$$f(x) = 2 \left(\sin x - \frac{1}{2} \sin 2x + \frac{1}{3} \sin 3x - \dots \right)$$

This equality holds for $-\pi < x < \pi$. At the points $\pm \pi$ the sum of the series turns into zero.

Note that the functions in Examples 2 and 3 coincide in the interval $[0, \pi]$ (they are equal to x within this interval). But since they are extended in different ways to the interval $[-\pi, 0)$ (the former as an even function

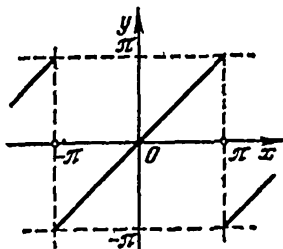


Fig. 258

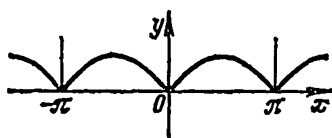


Fig. 259

and the latter as an odd function) their expansions into Fourier's series differ.

(4) Let us find the Fourier series of the function $f(x) = |\sin x|$ (see Fig. 259). Since the function is even we have $b_n = 0$ for all the values of n . Furthermore,

$$\begin{aligned} a_0 &= \frac{2}{\pi} \int_0^{\pi} \sin x \, dx = \frac{4}{\pi} \quad \text{and} \\ a_n &= \frac{2}{\pi} \int_0^{\pi} \sin x \cos nx \, dx = \frac{1}{\pi} \int_0^{\pi} [\sin(n+1)x - \sin(n-1)x] \, dx = \\ &= \begin{cases} 0, & \text{if } n \text{ is odd} \\ -\frac{4}{\pi(n^2-1)}, & \text{if } n \text{ is even} \end{cases} \end{aligned}$$

Consequently,

$$|\sin x| = \frac{2}{\pi} - \frac{4}{\pi} \left(\frac{1}{3} \cos 2x + \frac{1}{15} \cos 4x + \dots + \frac{1}{(2n)^2 - 1} \cos 2nx + \dots \right)$$

The translation of the origin turns the function $|\sin x|$ into $f(x) = |\cos x|$:

$$|\cos x| = \left| \sin \left(x - \frac{\pi}{2} \right) \right|$$

Therefore,

$$|\cos x| = \frac{2}{\pi} - \frac{4}{\pi} \left[\frac{1}{3} \cos 2x \left(x - \frac{\pi}{2} \right) + \frac{1}{15} \cos 4x \left(x - \frac{\pi}{2} \right) + \dots \right] = \\ = \frac{2}{\pi} + \frac{4}{\pi} \left(\frac{1}{3} \cos 2x - \frac{1}{15} \cos 4x + \dots \right)$$

(5) Let us expand into Fourier's series the function defined as $f(x) = x^2$ for $-1 \leq x \leq 1$. Extending this function periodically to the whole Ox -axis we obtain a periodic function with period 2. Using the formulas of Sec. 201, II and taking into account that the resultant function is even we obtain

$$a_0 = 2 \int_0^1 x^2 dx = \frac{2}{3} \quad \text{and} \quad a_n = 2 \int_0^1 x^2 \cos n\pi x dx = \\ = \left[\frac{x^2 \sin n\pi x}{n\pi} + \frac{2x \cos n\pi x}{\pi^2 n^2} - \frac{2 \sin n\pi x}{\pi^3 n^3} \right]_0^1 = \frac{4}{\pi^2 n^2} \cos n\pi$$

whence

$$f(x) = \frac{1}{3} - \frac{4}{\pi^2} \left(\cos \pi x - \frac{1}{2^2} \cos 2\pi x + \frac{1}{3^2} \cos 3\pi x - \dots \right)$$

202*. **Fourier Series in Complex Form.** Using complex numbers and Euler's formulas we can represent the Fourier series of a given function $f(x)$ defined in the interval $[-\pi, \pi]$ as a series with complex terms. This new form of a Fourier series is convenient for many mathematical and practical applications and can be easily memorized.

Consider a Fourier series

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

Let us replace $\cos nx$ and $\sin nx$ by their expressions given by Euler's formulas:

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2}, \quad \sin nx = \frac{e^{inx} - e^{-inx}}{2i}$$

Then

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left(a_n \frac{e^{inx} + e^{-inx}}{2} + b_n \frac{e^{inx} - e^{-inx}}{2i} \right) = \\ = \frac{1}{2} a_0 + \frac{1}{2} \sum_{n=1}^{\infty} [e^{inx} (a_n - ib_n) + e^{-inx} (a_n + ib_n)] \quad (*)$$

Let us denote the complex number $a_n - ib_n$ by c_n :

$$c_n = a_n - ib_n$$

Using the formulas for a_n and b_n (see Sec. 200) we can write

$$c_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) (\cos nx - i \sin nx) dx = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx \quad (**)$$

The complex number $a_n + ib_n$ is the complex conjugate of c_n :

$$a_n + ib_n = \overline{c_n}$$

The expression of $\overline{c_n}$ is obtained from that of c_n by substituting $-n$ for n . In this connection we introduce the notation

$$c_{-n} = \overline{c_n} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e^{inx} dx \quad (***)$$

Since $a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx$ the expression of a_0 is obtained from formula (**) if we put $n=0$; therefore we can write $a_0 = c_0$.

Using the new notation we now write down series (*) in the form

$$f(x) = \frac{1}{2} c_0 + \frac{1}{2} \sum_{n=1}^{\infty} c_n e^{inx} + \frac{1}{2} \sum_{n=1}^{\infty} c_{-n} e^{-inx}$$

or, briefly,

$$f(x) = \frac{1}{2} \sum_{n=-\infty}^{\infty} c_n e^{inx} \quad \text{where} \quad c_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx$$

In the latter formula the index of summation n runs through all the integral values from $-\infty$ to $+\infty$. Since the complex number e^{-inx} is the complex conjugate of e^{inx} the terms of the series with indices n and $-n$ are complex expressions of the form

$$\frac{1}{2} c_n e^{inx} \quad \text{and} \quad \frac{1}{2} c_{-n} e^{-inx} = \frac{1}{2} \overline{c_n e^{inx}}$$

Their sum

$$\begin{aligned} \frac{1}{2} (c_n e^{inx} + \overline{c_n e^{inx}}) &= \operatorname{Re} [(a_n - ib_n) (\cos nx + i \sin nx)] = \\ &= a_n \cos nx + b_n \sin nx \end{aligned}$$

is a real function representing the n th harmonic. It can be rewritten (see pp. 726, 727) as

$$a_n \cos nx + b_n \sin nx = A_n \cos (nx - \varphi_n)$$

where $A_n = \sqrt{a_n^2 + b_n^2} = |c_n|$ is the amplitude; since $\tan \varphi_n = \frac{b_n}{a_n}$, the initial phase is expressed as $\varphi_n = \arg \overline{c_n} = -\arg c_n$.

The sequence of complex numbers c_n is called the *spectral sequence* of the function $f(x)$; the real sequence $|c_n|$ is termed the

amplitude spectrum of $f(x)$ and the sequence $\varphi_n = -\arg c_n$ the phase spectrum. It is readily seen that $|c_{-n}| = |c_n|$ and $\varphi_{-n} = -\varphi_n$.

Amplitude and phase spectra play an important role in the study of periodic processes when a compound vibration is represented as a sum of simple harmonics, especially, in electrical engineering. In the general case we deal with a function $f(x)$ defined in an interval $(-l, l)$ for which the Fourier trigonometric expansion is constructed. In this case Fourier's series is written in complex form as

$$f(x) = \frac{1}{2} \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n x} \quad \text{where } c_n = \frac{1}{l} \int_{-l}^l f(x) e^{-i\omega_n x} dx$$

$$\text{and } \omega_n = \frac{\pi n}{l}$$

Let the reader derive this formula.

Using the integration rule for complex functions (see Sec. 93) we can apply formula (**) to computing Fourier's coefficients. In these calculations one should take into account that

$$e^{ln\pi} = e^{-ln\pi} = \cos n\pi = (-1)^n$$

Example. Let us construct the Fourier series of the function $f(x) = e^x$ considered in the interval $[-\pi, \pi]$. We have

$$c_n = \frac{1}{\pi} \int_{-\pi}^{\pi} e^x e^{-inx} dx = \frac{1}{\pi} \int_{-\pi}^{\pi} e^{x(1-in)} dx = \frac{1}{\pi} \frac{e^{x(1-in)}}{1-in} \Big|_{-\pi}^{\pi} =$$

$$= \frac{1}{\pi(1-in)} (e^{\pi} e^{-ln\pi} - e^{-\pi} e^{ln\pi}) = \frac{(-1)^n}{\pi(1-in)} (e^{\pi} - e^{-\pi})$$

Consequently,

$$f(x) = \frac{e^{\pi} - e^{-\pi}}{2\pi} \sum_{n=-\infty}^{\infty} (-1)^n \frac{e^{inx}}{1-in}$$

As was mentioned, to obtain the ordinary (real) form of the Fourier series we should combine the terms with indices n and $-n$:

$$(-1)^n \frac{e^{inx}}{1-in} + (-1)^{-n} \frac{e^{-inx}}{1+in} = (-1)^n \frac{e^{inx} + e^{-inx} + ni(e^{inx} - e^{-inx})}{1+n^2} =$$

$$= (-1)^n \frac{2 \cos nx - 2n \sin nx}{1+n^2}$$

This results in

$$f(x) = \frac{e^{\pi} - e^{-\pi}}{\pi} \left[\frac{1}{2} + \sum_{n=1}^{\infty} (-1)^n \left(\frac{\cos nx}{1+n^2} - \frac{n \sin nx}{1+n^2} \right) \right]$$

203. Practical Harmonic Analysis. If a function is specified by a simple analytical formula its Fourier coefficients can usually be found by means of integration, which leads to an explicit expres-

sion for its Fourier series. But there are many practical cases when the precise computation of the integrals expressing Fourier's coefficients encounters considerable difficulties or is impossible; this is also the case when the function is specified by its graph or by a table. Experimental data are often compiled into a table or are represented graphically by means of a point-by-point plotting or with the aid of a self-recording apparatus. Then there arises the problem of finding an approximate analytical formula for the function in question. For this purpose we can use trigonometric series provided that it is known that the function admits of a sufficiently accurate approximation by the sum of the first several terms of its Fourier series. The problem reduces to the computation of Fourier's coefficients; to solve the problem we can apply one of the approximate methods of computation of integrals.

The division of applied mathematics dealing with trigonometric approximations to functions specified graphically or by a table is termed *practical harmonic* or *wave* (or *Fourier*) *analysis*. There are various techniques simplifying the computation of Fourier's coefficients and providing standard computational schemes which take into account the peculiarities of the integrals in question. In this section we shall briefly discuss one of these methods using special "stencils".

Suppose we are given a function $y=f(x)$ defined in the interval $[0, 2\pi]$. Let us assume that, irrespective of the way the function is represented originally, its graph is known. We shall also suppose that, if necessary, the coordinate system is given a parallel displacement so that the whole graph of the function lies above the Ox -axis as close to it as possible (see Fig. 260). Such a displacement only affects the value of the constant term a_0 in the Fourier series, which is inessential since the necessary correction can easily be introduced; at the same time this makes it possible to avoid negative and too large positive values of the function.

To construct an approximating Fourier polynomial for the given function it is necessary to compute the first several Fourier coefficients

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx$$

and

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx$$

These integrals are evaluated by means of one of the numerical integration formulas; usually we apply the rectangle formula which is the simplest. The interval $[0, 2\pi]$ is divided into n equal parts

by means of the equidistant points of division $x_0=0, x_1, x_2, \dots, x_{n-1}, x_n=2\pi$ where $x_i = i \frac{2\pi}{n}$, $i=0, 1, 2, \dots, n$, the step of the calculations being $\Delta x = \frac{2\pi}{n}$. The expressions for the Fourier coefficients are then approximately written as

$$a_k \approx \frac{2}{n} \sum_{i=0}^{n-1} y_i \cos kx_i, \quad b_k \approx \frac{2}{n} \sum_{i=0}^{n-1} y_i \sin kx_i \quad (*)$$

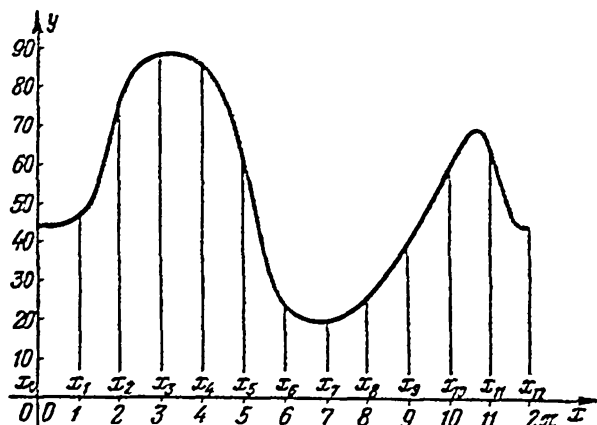


Fig. 260

where $y_i = f(x_i)$. Taking into account the peculiarities of the factors $\cos kx_i$ and $\sin kx_i$ we usually choose $n=12$ or $n=24$ or, if a higher accuracy is needed, $n=48$. In this section we shall take $n=12$. The reader can easily verify that in this case each of the 12 values of the function involved in formulas (*) is multiplied by one of the numbers

$$\begin{aligned} \cos 0 = \sin \frac{\pi}{2} = 1, \quad \cos \frac{\pi}{6} = \sin \frac{\pi}{3} = 0.87, \\ \cos \frac{\pi}{3} = \sin \frac{\pi}{6} = 0.50, \quad \cos \frac{\pi}{2} = \sin 0 = 0 \end{aligned}$$

taken with $+$ or $-$.

We shall discuss a convenient computation schedule known as the *12-ordinate scheme*.

The computations begin with compiling a table of 4 columns and 12 lines. The first column contains the successive numbers of the points of division $x_0, x_1, \dots, x_{11}$ of the interval $[0, 2\pi]$ and the second column contains the values of the ordinates corresponding to these points (if the function is specified graphically the ordinates are directly read off from the graph). When filling this column we should choose a sufficiently short scale along the y -axis

so that the ordinates are conveniently expressed by integers. In the third column we write the values of the products of the corresponding ordinates by $\cos 30^\circ = 0.87$ and in the fourth column their products by $\cos 60^\circ = 0.50$ (except the lines numbered 0, 3, 6 and 9 whose entries are crossed out since they correspond to the products of the ordinates by the cosines of arcs integer multiple of π or $\frac{\pi}{2}$).

After the table has been compiled (it is composed for every given function) we proceed to the computation of Fourier's coefficients. The coefficient a_0 is found directly by adding together the numbers in the second column and dividing the result by 6. To compute the coefficient a_k or b_k , $k > 0$, we use a standard stencil which is a replica of the table (but without numbers) made of a transparent material for each of the Fourier coefficients a_1 , b_1 , ..., a_6 , b_6 .

The positions corresponding to positive summands in formulas (*) and those corresponding to negative summands are marked differently so that it is possible to distinguish between them (for instance, the former can be outlined with paint or heavy line and the latter with some other paint or thin line).

0	44	—	—
1	46	40	23
2	76	66	38
3	88	—	—
4	86	75	43
5	63	55	36.5
6	24	—	—
7	20	17	10
8	26	22	13
9	40	—	—
10	58	50	29
11	65	51	32.5

a_1				b_1				a_2				b_2			
0				0				0				0			
1				1				1				1			
2				2				2				2			
3				3				3				3			
4				4				4				4			
5				5				5				5			
6				6				6				6			
7				7				7				7			
8				8				8				8			
9				9				9				9			
10				10				10				10			
11				11				11				11			

Fig. 261

In Fig. 261 we see the stencils for the first four coefficients a_1 , b_1 , a_2 and b_2 (for the 12-ordinate scheme). Let the reader examine their structure using formulas (*).

Applying the stencil for the coefficient a_k or b_k to the table we choose from the table the numbers corresponding to positive and negative summands in formulas (*). On computing the sums of the numbers in these groups and subtracting the latter sum from the former we obtain $6a_k$ or $6b_k$, and to find a_k or b_k it only remains to divide the result by 6. Carrying out these calculations for the table on p. 745 we obtain

$$\begin{aligned} a_0 &= \frac{44+46+76+88+86+63+24+20+26+40+58+65}{6} = \frac{636}{6} = 106 \\ a_1 &= \frac{(44+40+38+29+51)-(43+55+24+17+13)}{6} = \frac{50}{6} \approx 8.3 \\ b_1 &= \frac{(23+66+88+75+36.5)-(10+22+40+50+32.5)}{6} = \frac{134}{6} \approx 22.3 \\ a_2 &= \frac{(44+23+36.5+24+10+32.5)-(38+88+43+13+40+29)}{6} = -\frac{81}{6} \approx -13.5 \\ b_2 &= \frac{(40+66+17+22)-(75+55+50+51)}{6} = -\frac{86}{6} \approx -14.3 \end{aligned}$$

Hence, we get the following approximate representation of the tabulated function $f(x)$ as a trigonometric polynomial:

$$f(x) \approx 53 + (8.3 \cos x + 22.3 \sin x) - (13.5 \cos 2x + 14.3 \sin 2x)$$

For more detail concerning practical harmonic analysis we refer the reader to [8] and [13].

§ 2. Some Further Topics in the Theory of Fourier Series

204*. Orthogonal Systems of Functions. Here we shall briefly discuss a more general problem than that of expanding a function into a trigonometric series. Namely, we shall be concerned with the problem of expanding functions in series with respect to a given *orthogonal system of functions*. The functions $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$, ... we shall deal with are supposed to be defined and continuous in an interval $[a, b]$.

Definition. A system of functions $\varphi_1(x)$, $\varphi_2(x)$, ..., $\varphi_n(x)$, ... is said to be *orthogonal* in the interval $[a, b]$ if the integral of the product of any two different functions of the system taken over the interval $[a, b]$ is equal to zero:

$$\int_a^b \varphi_n(x) \varphi_m(x) dx = 0, \quad n \neq m$$

Let k_n denote the integral of the square of the function $\varphi_n(x)$:

$$k_n = \int_a^b \varphi_n^2(x) dx$$

We shall suppose that all the numbers k_n , $n=1, 2, \dots$, are positive, i.e. there are no functions identically equal to zero among the functions $\varphi_n(x)$.

If we multiply each function $\varphi_n(x)$ by the factor $\frac{1}{\sqrt{k_n}}$ the new functions

$$\psi_1(x) = \frac{1}{\sqrt{k_1}} \varphi_1(x), \psi_2(x) = \frac{1}{\sqrt{k_2}} \varphi_2(x), \dots, \psi_n(x) = \frac{1}{\sqrt{k_n}} \varphi_n(x), \dots$$

of course again form an orthogonal system. The integral of the square of each new function is equal to unity:

$$\int_a^b \psi_n^2(x) dx = \frac{1}{k_n} \int_a^b \varphi_n^2(x) dx = 1$$

The functions $\psi_1(x), \dots, \psi_n(x), \dots$ thus satisfy the conditions

$$\int_a^b \psi_m(x) \psi_n(x) dx = \begin{cases} 0 & \text{for } m \neq n \\ 1 & \text{for } m = n \end{cases} \quad (*)$$

A functional system of this kind is termed an *orthogonal system of normalized functions* or, briefly, an *orthonormal system*.

Examples. (1) The system of functions $1, \cos x, \sin x, \dots, \cos nx, \sin nx, \dots$ is an orthogonal system for the interval $[-\pi, \pi]$, which follows from formulas (1)-(5) of Sec. 199. These formulas also imply that the system

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin nx}{\sqrt{\pi}}, \dots$$

is orthonormal.

(2) The functions $\sin x, \sin 2x, \dots, \sin nx, \dots$ form an orthogonal system in the interval $[0, \pi]$.

Indeed, if $m \neq n$ we have

$$\int_0^\pi \sin mx \sin nx dx = \left[\frac{\sin(m-n)x}{2(m-n)} - \frac{\sin(m+n)x}{2(m+n)} \right] \Big|_0^\pi = 0$$

(see Formula 62 in the table of integrals at the end of the book). Since

$$\int_0^\pi \sin^2 nx dx = \left[\frac{x}{2} - \frac{\sin 2nx}{2n} \right] \Big|_0^\pi = \frac{\pi}{2}$$

this system is normalized if all the functions are divided by $\sqrt{\frac{\pi}{2}}$.

(3) The system of functions $1, \cos x, \cos 2x, \dots, \cos nx, \dots$ considered in the interval $[0, \pi]$ is orthogonal. Let the reader prove this.

Let $f(x)$ be a function defined in an interval $[a, b]$. Suppose that it can be represented in this interval as the sum of a series in functions $\psi_k(x)$ ($k=1, 2, \dots$) with constant factors c_k where $\psi_1(x), \psi_2(x), \dots, \psi_n(x), \dots$ is an orthonormal system of functions in this interval:

$$f(x) = c_1\psi_1(x) + c_2\psi_2(x) + \dots + c_n\psi_n(x) + \dots \quad (**)$$

To determine the *coefficients of the expansion*, that is the numbers c_k , we multiply, in succession, both sides of equality (**) by the functions $\psi_1(x), \psi_2(x), \psi_3(x), \dots, \psi_n(x), \dots$ and integrate the results over the interval $[a, b]$. We of course assume that the term-by-term integration of the series is permissible, and thus, by virtue of relations (*), obtain

$$\int_a^b f(x) \psi_n(x) dx = \sum_{k=1}^{\infty} c_k \int_a^b \psi_k(x) \psi_n(x) dx = c_n$$

The coefficients c_n specified by these formulas are called the (*generalized*) *Fourier coefficients of the function $f(x)$ with respect to the given orthonormal system*.

In particular, if the functions $\psi_n(x)$ are the trigonometric functions indicated in Example 1, series (**) is nothing but the Fourier (trigonometric) series of the function $f(x)$.

Expansion of form (**) is called the (*generalized*) *Fourier series* of the function $f(x)$. Such expansions are widely used in the solutions of many important problems of mathematical physics; a number of such problems are discussed in [4].

205. Mean Square Deviation. Minimizing Property of Fourier Coefficients.

I. In our course we many times dealt with the problem of replacing a given function by an approximating function of a simpler nature. For instance, if a function can be expanded into Taylor's series in an interval it can be approximated with a Taylor polynomial, the accuracy of the approximation usually increasing together with the degree n of the polynomial. As was mentioned in Sec. 200, Fourier's polynomials can also be used for this purpose.

If a function $f(x)$ defined in an interval $[a, b]$ is replaced by another function $\varphi(x)$ the absolute value of the error, i.e. the quantity $|f(x) - \varphi(x)|$, may assume different values at different

points. For example, this is clearly seen in Fig. 249 (p. 709) showing the graphs of the function $\sin x$ and of its approximating Taylor polynomial or in Fig. 257 (p. 738) where the approximation of the function $|x|$ with Fourier polynomials is demonstrated.

There are situations when it is convenient to estimate the accuracy of an approximation for the whole interval in which the function $f(x)$ is considered. Let the function $f(x)$ and the approximating function $\varphi(x)$ be continuous throughout the closed interval $[a, b]$. Then, according to the property of continuous functions, there exists the greatest value of the modulus of the difference of these functions which we denote $\rho\{f, \varphi\}$:

$$\rho\{f, \varphi\} = \max_{x \in [a, b]} |f(x) - \varphi(x)|$$

The absolute value of the error of the approximation of $f(x)$ with $\varphi(x)$ satisfies the inequality

$$|f(x) - \varphi(x)| \leq \rho\{f, \varphi\}$$

at each point of the interval $[a, b]$.

The magnitude $\rho\{f, \varphi\}$ is called the *uniform deviation* of the function $\varphi(x)$ from the function $f(x)$.

P.L. Chebyshev¹ posed and solved the problem of determining, among all the polynomials of the n th degree, a polynomial $P_n(x)$ such that its deviation from the given function $f(x)$ in the given interval $[a, b]$ attains the least possible value. This means that

$$\rho\{f, P_n\} = \max_{x \in [a, b]} |f(x) - P_n(x)| \leq \rho\{f, Q_n\}$$

where $Q_n(x)$ is any other polynomial of the n th degree.

Such polynomials are widely used in various divisions of mathematics and its applications (e.g. in the theory of mechanisms). For further detail on Chebyshev's polynomials we refer the reader to [8] and [13].

There are also problems of another kind in which the deviation of the function $\varphi(x)$ from the function $f(x)$ at separate points is inessential and when we are only interested in its characteristic "in the mean". As a measure $r\{f, \varphi\}$ of the deviation of this kind the value of the integral of the square of the difference between the functions $f(x)$ and $\varphi(x)$ is usually taken:

$$r\{f, \varphi\} = \int_a^b [f(x) - \varphi(x)]^2 dx$$

¹ P.L. Chebyshev (1821-1894), a prominent Russian mathematician. His studies affected greatly the development of mathematics. Their feature was a close connection between mathematical questions and problems of natural and engineering sciences.

The magnitude $\sqrt{\frac{1}{b-a} r\{f, \varphi\}}$ is called the *mean square deviation*¹. It should be noted that even for small values of the magnitude $r\{f, \varphi\}$ the difference between the values of the functions $f(x)$ and $\varphi(x)$ can be arbitrarily large at some separate points.

II. Let us come back to the subject of Sec. 204. Suppose that we are given an orthonormal system of functions in an interval $[a, b]$: $\psi_1(x), \psi_2(x), \dots, \psi_n(x), \dots$. Let $f(x)$ be a piecewise smooth function defined in the same interval $[a, b]$. Now we fix a natural number n and pose the following problem:

Among all linear combinations

$$\sigma_n(x) = \gamma_1 \psi_1(x) + \gamma_2 \psi_2(x) + \dots + \gamma_n \psi_n(x) \quad (*)$$

of the first n functions $\psi_k(x)$ ($k=1, 2, \dots, n$) with arbitrary real coefficients γ_k , it is required to find the one for which the mean square deviation of that linear combination from the function $f(x)$ achieves the least possible value.

In other words, the coefficients γ_k should be chosen so that they minimize the integral

$$r\{f, \sigma_n\} = \int_a^b [f(x) - \sigma_n(x)]^2 dx$$

Let us represent the magnitude $r\{f, \sigma_n\}$ in the form

$$r\{f, \sigma_n\} = \int_a^b f^2(x) dx - 2 \int_a^b f(x) \sigma_n(x) dx + \int_a^b \sigma_n^2(x) dx$$

and replace $\sigma_n(x)$ by its expression. When $\sigma_n(x)$ is squared the integrals of the pairwise products of different functions $\psi_k(x)$ turn into zero while the integrals of the squares of the functions are equal to unity (see formulas $(*)$ in Sec. 204). Since

$\int_a^b f(x) \psi_k(x) dx = c_k$ is the Fourier coefficient of the function $f(x)$, with respect to the given orthonormal system, corresponding to the function $\psi_k(x)$ the expression of $r\{f, \sigma_n\}$ can be rewritten as

$$r\{f, \sigma_n\} = \int_a^b f^2(x) dx - 2 \sum_{k=1}^n \gamma_k c_k + \sum_{k=1}^n \gamma_k^2$$

Combining the first two sums and completing the squares by adding and subtracting the quantities c_k^2 ($k=1, 2, \dots, n$) we arrive

¹ The question discussed here is related to one of the most effective methods of solution of problems of numerical analysis known as the *least square method* (e.g. see [12], [13]).

at the final formula

$$r\{f, \sigma_n\} = \int_a^b f^2(x) dx - \sum_{k=1}^n c_k^2 + \sum_{k=1}^n (\gamma_k - c_k)^2 \quad (**)$$

The first two terms on the right-hand side of formula (**) do not involve the numbers γ_k , and therefore the magnitude $r\{f, \sigma_n\}$ attains its minimum value when the last sum on the right turns into zero. This is achieved for

$$\gamma_1 = c_1, \quad \gamma_2 = c_2, \quad \dots, \quad \gamma_n = c_n$$

Thus, among all linear combinations (*), the one coinciding with the n th partial sum of the generalized Fourier series of the function $f(x)$ gives the magnitude $r\{f, \sigma_n\}$ its least possible value. This is the so-called *minimizing property* of Fourier's coefficients.

Thus, let

$$s_n(x) = c_1\psi_1(x) + c_2\psi_2(x) + \dots + c_n\psi_n(x)$$

where c_n are the Fourier coefficients of the function $f(x)$. Equality (**) implies that

$$r\{f, s_n\} = \int_a^b [f(x) - s_n(x)]^2 dx = \int_a^b f^2(x) dx - \sum_{k=1}^n c_k^2 \quad (***)$$

Since the integral of the square of a function is always non-negative we conclude that for any n there holds the relation

$$\sum_{k=1}^n c_k^2 \leq \int_a^b f^2(x) dx$$

known as *Bessel's¹ inequality*. It shows that the partial sums of the series $\sum_{k=1}^{\infty} c_k^2$ have a common upper bound equal to

the integral $\int_a^b f^2(x) dx$; hence, this numerical series is convergent.

By the way, this means that the numbers c_k (the Fourier coefficients of the function $f(x)$) tend to zero as $k \rightarrow \infty$.

206. Convergence in the Mean. Parseval's² Relation. Lyapunov's³ Theorem. Let us come back to formula (***) of Sec. 205. It is clear that if the number n is increased, that is the linear combinations are taken with greater number of functions $\psi_k(x)$, the

¹ F. W. Bessel (1784-1846), a German astronomer and mathematician.

² M. A. Parseval (died 1836), a French mathematician.

³ A. M. Lyapunov (1857-1918), a noted Russian mathematician. He founded the theory of stability of solutions of differential equations and also contributed many important results to the theory of probability and to the theory of equations of mathematical physics.

mean square deviation can only decrease. It is important for applications that the minimizing coefficients for a fixed value of n remain among the minimizing coefficients for greater values of n , and that only new coefficients are added to them.

In this connection it appears natural to pose the question as to whether or not the mean square deviation tends to zero as n is increased indefinitely, i.e. whether or not $r \{f, s_n\} \rightarrow 0$ for the given function $f(x)$. If the answer to the question is affirmative we have

$$\lim_{n \rightarrow \infty} \int_a^b [f(x) - s_n(x)]^2 dx = 0$$

In this case we say that the sequence $s_n(x)$ formed of the partial sums of Fourier's series is *convergent in the mean* to the function $f(x)$ or that the Fourier series of the function $f(x)$ *converges to it in the mean*. The convergence in the

mean implies that the numerical series $\sum_{k=1}^{\infty} c_k^2$ is convergent

and that its sum is equal to $\int_a^b f^2(x) dx$:

$$\sum_{k=1}^{\infty} c_k^2 = \int_a^b f^2(x) dx$$

The latter equality is called *Parseval's relation*.

III. Applying the general results discussed here to Fourier's series with respect to the trigonometric system we can state the following theorem whose rigorous proof was given by A. M. Lyapunov:

Lyapunov's Theorem. Let the function $f(x)$ be defined and piecewise continuous in the interval $[-\pi, \pi]$. Then its Fourier polynomials

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx)$$

are convergent in the mean to the function $f(x)$, that is

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} [f(x) - S_n(x)]^2 dx = 0$$

Besides, there holds the formula

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(x) dx$$

This theorem shows that any function of this type and, in particular, every continuous function, can be approximated in the mean with an arbitrary accuracy by its Fourier polynomials. Lyapunov's theorem has many important applications. It also applies to a wider class of functions, namely, to functions integrable together with their squares over the interval $[-\pi, \pi]$ (at least in the improper sense). These are the so-called *square-integrable* (*square-summable*) functions.

§ 3*. Fourier Integral

207*. Fourier Integral. As was shown in Sec. 201, II, the expansion of a function $f(x)$ defined in an interval $[-l, l]$ into Fourier's series has the form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos \omega_n x + b_n \sin \omega_n x)$$

where

$$\omega_n = \frac{\pi n}{l}, \quad a_n = \frac{1}{l} \int_{-l}^l f(t) \cos \omega_n t \, dt \quad \text{and} \quad b_n = \frac{1}{l} \int_{-l}^l f(t) \sin \omega_n t \, dt$$

(For the sake of convenience we denoted by t the variable of integration in the integrals specifying the coefficients a_n and b_n .)

The substitution of the expressions of a_n and b_n into the Fourier series leads to

$$\begin{aligned} f(x) &= \frac{1}{2l} \int_{-l}^l f(t) \, dt + \\ &+ \frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) [\cos \omega_n t \cos \omega_n x + \sin \omega_n t \sin \omega_n x] \, dt = \\ &= \frac{1}{2l} \int_{-l}^l f(t) \, dt + \frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \omega_n (t-x) \, dt \end{aligned} \quad (*)$$

Let us suppose that the function $f(x)$ is defined throughout the x -axis and that it is piecewise smooth in any finite interval $[-l, l]$ and thus can be expanded into Fourier's series in every such interval. Besides, we shall assume that the improper integral of the absolute value of this function taken from $-\infty$ to $+\infty$ is convergent:

$$\int_{-\infty}^{\infty} |f(t)| \, dt = Q$$

where Q is a finite number (such functions are said to be *absolutely integrable* or *absolutely summable*).

Let us first consider the function $f(x)$ in the interval $[-l, l]$ and expand it into the corresponding Fourier series which we write in form (*). Then we make l increase indefinitely and pass to the limiting form of formula (*). The first summand tends to zero as $l \rightarrow \infty$:

$$\left| \frac{1}{2l} \int_{-l}^l f(t) dt \right| \leq \frac{1}{2l} \int_{-l}^l |f(t)| dt \leq \frac{Q}{2l} \rightarrow 0$$

(Here we use the inequality connecting the absolute value of the integral of $f(x)$ and the integral of $|f(x)|$ proved in Sec. 90.)

Let us introduce the notation

$$\Delta\omega = \omega_{n+1} - \omega_n = \frac{\pi}{l} \quad (n=0, 1, 2, \dots)$$

Then the remaining sum can be rewritten in the form

$$\frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \omega_n(t-x) dt = \frac{1}{\pi} \sum_{n=1}^{\infty} \Delta\omega \int_{-l}^l f(t) \cos \omega_n(t-x) dt$$

Note that the integral $\varphi(\omega) = \int_{-\infty}^{\infty} f(t) \cos \omega(t-x) dt$ is convergent since the absolute value of the integrand does not exceed the function $|f(t)|$ which is integrable over $(-\infty, \infty)$. Moreover, as function of the parameter ω this integral is regularly convergent, and hence, if $f(x)$ is a continuous function, the function $\varphi(\omega)$ is continuous as well (see Sec. 137). For sufficiently large values of l the integral $\int_{-l}^l f(t) \cos \omega_n(t-x) dt$ can be regarded as being almost equal to the value of the function $\varphi(\omega)$ assumed for $\omega = \omega_n$, that is,

$$\int_{-l}^l f(t) \cos \omega_n(t-x) dt \approx \varphi(\omega_n)$$

The above sum can therefore be rewritten as

$$\frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \omega_n(t-x) dt = \frac{1}{\pi} \sum_{n=1}^{\infty} \varphi(\omega_n) \Delta\omega$$

In the sum $\sum_{n=1}^{\infty} \varphi(\omega_n) \Delta\omega$ the quantity $\Delta\omega = \frac{\pi}{l}$ tends to zero as $l \rightarrow \infty$ while, in the limit, the variable ω_n runs through all the values of ω from 0 to ∞ . Therefore it is natural to expect that

this sum tends to the integral

$$\int_0^{\infty} \varphi(\omega) d\omega = \int_0^{\infty} d\omega \int_{-\infty}^{\infty} f(t) \cos \omega(t-x) dt$$

On substituting the expressions obtained into the formula of $f(x)$ we obtain

$$f(x) = \frac{1}{\pi} \int_0^{\infty} d\omega \int_{-\infty}^{\infty} f(t) \cos \omega(t-x) dt \quad (**)$$

The integral on the right is called *Fourier's integral* and the formula itself is *Fourier's integral formula*.

The above argument is not a rigorous proof; the latter involves some more sophisticated considerations which we do not present here. We only note that formula (**) holds for all the points of continuity of the function $f(x)$. At every point of discontinuity x_0 Fourier's integral (like the sum of Fourier's series) assumes the value $\frac{f(x_0-0) + f(x_0+0)}{2}$.

Since $\cos \omega(t-x) = \cos \omega t \cos \omega x + \sin \omega t \sin \omega x$ formula (**) can be rewritten in the form

$$\begin{aligned} f(x) = & \frac{1}{\pi} \int_0^{\infty} \cos \omega x d\omega \int_{-\infty}^{\infty} f(t) \cos \omega t dt + \\ & + \frac{1}{\pi} \int_0^{\infty} \sin \omega x d\omega \int_{-\infty}^{\infty} f(t) \sin \omega t dt \end{aligned} \quad (***)$$

Denoting the inner integrals $A(\omega)$ and $B(\omega)$, i.e.

$$A(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos \omega t dt \quad \text{and} \quad B(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \sin \omega t dt$$

we receive the relation

$$f(x) = \int_0^{\infty} [A(\omega) \cos \omega x + B(\omega) \sin \omega x] d\omega$$

For a periodic function with period $2l$ we have its Fourier expansion consisting of an infinite number of separate (discrete) harmonics with frequencies $\omega_n = \frac{\pi n}{l}$ differing from each other by the constant quantity $\Delta\omega = \frac{\pi}{l}$. Every such harmonic has a definite amplitude. As we say, a periodic function has a *discrete spectrum* (which exactly means that it is representable

where Q is a finite number (such functions are said to be *absolutely integrable* or *absolutely summable*).

Let us first consider the function $f(x)$ in the interval $[-l, l]$ and expand it into the corresponding Fourier series which we write in form (*). Then we make l increase indefinitely and pass to the limiting form of formula (*). The first summand tends to zero as $l \rightarrow \infty$:

$$\left| \frac{1}{2l} \int_{-l}^l f(t) dt \right| \leq \frac{1}{2l} \int_{-l}^l |f(t)| dt \leq \frac{Q}{2l} \rightarrow 0$$

(Here we use the inequality connecting the absolute value of the integral of $f(x)$ and the integral of $|f(x)|$ proved in § 136.)

Let us introduce the notation

$$\Delta\omega = \omega_{n+1} - \omega_n = \frac{\pi}{l} \quad (n = 0, 1, 2, \dots)$$

Then the remaining sum can be rewritten in the form

$$\frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \omega_n(t-x) dt = \frac{1}{\pi} \sum_{n=1}^{\infty} \Delta\omega \int_{-l}^l f(t) \cos \omega_n(t-x) dt$$

Note that the integral $\varphi(\omega) = \int_{-\infty}^{\infty} f(t) \cos \omega(t-x) dt$ is continuous since the absolute value of the integrand does not exceed $|f(t)|$ which is integrable over $(-\infty, \infty)$ as function of the parameter ω this integral is regular, and hence, if $f(x)$ is a continuous function $\varphi(\omega)$ is continuous as well (see Sec. 137). For small

values of l the integral $\int_{-l}^l f(t) \cos \omega_n(t-x) dt$ can be almost equal to the value of the function $\varphi(\omega)$ for $\omega = \omega_n$, that is,

$$\int_{-l}^l f(t) \cos \omega_n(t-x) dt \approx \varphi(\omega_n)$$

The above sum can therefore be rewritten as

$$\frac{1}{l} \sum_{n=1}^{\infty} \int_{-l}^l f(t) \cos \omega_n(t-x) dt = \frac{1}{\pi} \sum_{n=1}^{\infty} \varphi(\omega_n) \Delta\omega$$

In the sum $\sum_{n=1}^{\infty} \varphi(\omega_n) \Delta\omega$ the quantity $\Delta\omega$ tends to zero as $l \rightarrow \infty$ while, in the limit, the variable ω_n takes on all values of ω from 0 to ∞ . Therefore it is

In the former case we have

$$f(x) = \frac{2}{\pi} \int_0^{\infty} \cos \omega x d\omega \int_0^1 \cos \omega t dt = \frac{2}{\pi} \int_0^{\infty} \frac{\cos \omega x \sin \omega}{\omega} d\omega$$

By the way, this result shows that

$$\frac{2}{\pi} \int_0^{\infty} \frac{\cos \omega x \sin \omega}{\omega} d\omega = \begin{cases} 1 & \text{for } -1 < x < 1 \\ \frac{1}{2} & \text{for } x = \pm 1 \\ 0 & \text{for } x < -1 \text{ and } x > 1 \end{cases}$$

In the latter case (the odd extension) we obtain

$$f(x) = \frac{2}{\pi} \int_0^{\infty} \sin \omega x d\omega \int_0^1 \sin \omega t dt = \frac{2}{\pi} \int_0^{\infty} \frac{\sin \omega x (1 - \cos \omega)}{\omega} d\omega$$

If we take half the sum of Fourier's integrals written above we obtain Fourier's integral formula for the function equal to $f(x)$ for $x > 0$ and to zero for $x < 0$ (Fig. 262): if

$$F(x) = \begin{cases} f(x) & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

then

$$F(x) = \frac{1}{\pi} \int_0^{\infty} \left(\cos \omega x \frac{\sin \omega}{\omega} + \sin \omega x \frac{1 - \cos \omega}{\omega} \right) d\omega$$

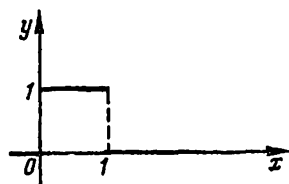


Fig. 262

(2) Let us construct Fourier's integrals for the function

$$f(x) = e^{-\beta x}, \quad \beta > 0, \quad x > 0$$

extended, as in Example 1, in both ways, to the negative half-axis Ox . For the even extension, using Formula 74 of the table of integrals at the end of the book, we receive

$$f(x) = \frac{2}{\pi} \int_0^{\infty} \cos \omega x d\omega \int_0^{\infty} e^{-\beta t} \cos \omega t dt = \frac{2}{\pi} \int_0^{\infty} \frac{\omega}{\omega^2 + \beta^2} \cos \omega x d\omega$$

For the odd extension we use Formula 73 of this table to obtain

$$f(x) = \frac{2}{\pi} \int_0^{\infty} \sin \omega x d\omega \int_0^{\infty} e^{-\beta t} \sin \omega t dt = \frac{2}{\pi} \int_0^{\infty} \frac{\omega}{\omega^2 + \beta^2} \sin \omega x d\omega$$

Since $f(0) = 1$ the point $x = 0$ is a point of discontinuity of the function $f(x)$ extended as an odd function. Therefore the corresponding Fourier integral is equal to zero but not to $f(0) = 1$

for $x=0$, that is, to the half-sum of the right-hand and left-hand limits of the function at that point. Finally, the function

$$\varphi(x) = \begin{cases} e^{-\beta x} & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

is representable in the form

$$\varphi(x) = \frac{1}{\pi} \int_0^{\infty} \frac{\beta \cos \omega x + \omega \sin \omega x}{\omega^2 + \beta^2} d\omega$$

209*. **Fourier Integral in Complex Form.** Fourier Transformation. The complex form of Fourier's integral is completely analogous to that of Fourier's series. By formula (***) of Sec. 205, we have

$$f(x) = \int_0^{\infty} [A(\omega) \cos \omega x + B(\omega) \sin \omega x] d\omega$$

where

$$A(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \cos \omega t dt \quad \text{and} \quad B(\omega) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(t) \sin \omega t dt$$

Replacing $\cos \omega x$ and $\sin \omega x$ in the Fourier integral by their expression given by Euler's formulas we derive

$$\begin{aligned} f(x) &= \int_0^{\infty} \left[A(\omega) \frac{e^{i\omega x} + e^{-i\omega x}}{2} + B(\omega) \frac{e^{i\omega x} - e^{-i\omega x}}{2i} \right] d\omega = \\ &= \frac{1}{2} \int_0^{\infty} \{ [A(\omega) - iB(\omega)] e^{i\omega x} + [A(\omega) + iB(\omega)] e^{-i\omega x} \} d\omega \end{aligned}$$

Let us put $\pi[A(\omega) - iB(\omega)] = F(\omega)$. Then, by formulas for $A(\omega)$ and $B(\omega)$, we can write

$$F(\omega) = \int_{-\infty}^{\infty} f(t) (\cos \omega t - i \sin \omega t) dt = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (*)$$

Furthermore, we have

$$\pi[A(\omega) + iB(\omega)] = \overline{F(\omega)} = \int_{-\infty}^{\infty} f(t) e^{i\omega t} dt$$

The function $F(\omega)$ is originally defined for $\omega > 0$. If we put, by definition, $F(-\omega) = \overline{F(\omega)}$, formula (*) will specify the extended function $F(\omega)$ both for the positive and negative values of ω . The substitution of the extended function $F(\omega)$ into Fourier's

integral now yields

$$f(x) = \frac{1}{2\pi} \int_0^{\infty} [F(\omega) e^{i\omega x} + F(-\omega) e^{-i\omega x}] d\omega$$

The integral of the second summand can be written as

$$\int_0^{\infty} F(-\omega) e^{-i\omega x} d\omega = \int_0^{-\infty} F(\omega) e^{i\omega x} d(-\omega) = \int_{-\infty}^0 F(\omega) e^{i\omega x} d\omega$$

Therefore, we lastly obtain

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega \quad (**)$$

where the function $F(\omega)$ is specified by the formula written above:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (*)$$

(The improper integral in formula (**)) is understood in the sense of Cauchy's *principal value*, i.e. as the limit $\lim_{l \rightarrow \infty} \int_{-l}^l F(\omega) e^{i\omega x} d\omega$.)

The function $F(\omega)$ is called the *Fourier transform* of the function $f(x)$ and, generally speaking, is a complex function of the real argument ω . The original function $f(x)$ is *Fourier's inverse transform* of the function $F(\omega)$.¹

Now suppose that the function $f(x)$ is originally defined in the interval $[0, \infty)$ and then extended as an *even function* and as an *odd function* to the interval $(-\infty, 0)$. Using the results of Sec. 207, we arrive at the following formulas:

$$f(x) = \frac{1}{\pi} \int_0^{\infty} F_c(\omega) \cos \omega x d\omega \quad \text{where} \quad F_c(\omega) = 2 \int_0^{\infty} f(t) \cos \omega t dt$$

¹ The number π is not involved symmetrically in formulas (*) and (**). We can represent the coefficient $\frac{1}{2\pi}$ as $\frac{1}{2\pi} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2\pi}}$ and attribute one of the factors $\frac{1}{\sqrt{2\pi}}$ to the function $F(\omega)$. Then both formulas will involve the same factor $\frac{1}{\sqrt{2\pi}}$.

for the former case and

$$f(x) = \frac{1}{\pi} \int_0^{\infty} F_s(\omega) \sin \omega x d\omega \quad \text{where} \quad F_s(\omega) = 2 \int_0^{\infty} f(t) \sin \omega t dt$$

for the latter case.

The functions $F_c(\omega)$ and $F_s(\omega)$ are called, respectively, *Fourier's cosine transform* and *Fourier's sine transform* of the function $f(x)$.

The passage from a given function to its Fourier transform is referred to as *Fourier's transformation*, the reverse passage being termed *Fourier's inverse transformation*. The theory of Fourier's transformation plays an important role in the solution of many problems of the theory of equations of mathematical physics. There are extensive tables in which various pairs of mutually corresponding functions $f(x)$ and $F(\omega)$ are given (e.g. see [14]).

Let us compare the formulas

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega, \quad F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

expressing the Fourier transformation with the complex form of Fourier's series (see Sec. 202). For a function $f(x)$ defined in an interval $[-l, l]$ we have

$$f(x) = \frac{1}{2} \sum_{n=-\infty}^{\infty} c_n e^{i\omega_n x}, \quad c_n = \frac{1}{l} \int_{-l}^l f(t) e^{-i\omega_n t} dt \quad \text{where} \quad \omega_n = \frac{\pi n}{l}$$

Thus, when we pass from a finite interval $[-l, l]$ to the infinite interval $(-\infty, \infty)$ the *spectral sequence* c_n is replaced by the Fourier transform $F(\omega)$ (which is also called the *spectral function* of $f(x)$) and the sum of the expressions $c_n e^{i\omega_n x}$ is replaced by the *integral* of the function $F(\omega) e^{i\omega x}$. The function $|F(\omega)|$ is termed the *amplitude spectrum* of the function $f(x)$ and $\varphi(\omega) = -\arg F(\omega)$ the *phase spectrum*.

Example. Let us find the Fourier transform, and the amplitude and phase spectra for the function

$$f(t) = \begin{cases} e^{-at} & \text{for } t > 0 \\ 0 & \text{for } t < 0, \end{cases} \quad (a > 0)$$

We have

$$F(\omega) = \int_0^{\infty} e^{-at} e^{-i\omega t} dt = -\frac{e^{-t(a+i\omega)}}{a+i\omega} \Big|_0^{\infty} = \frac{1}{a+i\omega}$$

Consequently,

$$|F(\omega)| = \frac{1}{\sqrt{a^2 + \omega^2}}, \quad \varphi(\omega) = -\arg F(\omega) = \arctan \frac{\omega}{a}$$

QUESTIONS

1. What is a trigonometric series? State the problem of expanding a function into Fourier's series.
2. Derive the formulas for the Fourier coefficients of a function $f(x)$ defined in the interval $[-\pi, \pi]$.
3. State the sufficient conditions for a function being representable by its Fourier series.
4. Enumerate the peculiarities of Fourier's expansions for even and odd functions.
5. Derive the formulas for the Fourier coefficients of a function $f(x)$ defined in an interval $[-l, l]$.
6. Write down the formulas for the Fourier coefficients of the expansions into Fourier's sine and cosine series of a function defined in the interval $[0, \pi]$.
- 7*. Write down the complex form of Fourier's series and the formulas for its coefficients.
8. Enumerate the methods for approximate computation of Fourier's coefficients.
- 9*. What do we call an orthogonal system of functions?
- 10*. State the problem of expanding a function into a series with respect to an orthogonal system of functions. How do we find the coefficients of such an expansion?
- 11*. What is the uniform deviation between two functions? What is the mean square deviation?
- 12*. Explain the minimizing property of the coefficients of Fourier's series.
- 13*. Prove Bessel's inequality.
- 14*. State the definition of convergence in the mean.
- 15*. Write down Parseval's relation.
- 16*. State Lyapunov's theorem.
- 17*. What is the Fourier integral?
- 18*. State sufficient conditions for a function being representable as Fourier's integral.
- 19*. How is Fourier's integral for even and for odd functions written?
- 20*. Write Fourier's integral in complex form.
- 21*. What is Fourier's transform of a function? What is Fourier's inverse transform?
- 22*. State the definitions of Fourier's cosine and sine transforms.
- 23*. What do we call the amplitude and the phase spectra of a function? What is the difference between the spectra of a function defined in a finite interval $[-l, l]$ and of a function defined in the interval $(-\infty, \infty)$?

TABLE OF INTEGRALS

In the formulas below a , b , m and n are constant parameters.

I. Integrals of Rational Functions

1. $\int (ax+b)^n dx = \frac{(ax+b)^{n+1}}{a(n+1)} + C$ where n is an integer different from -1
2. $\int \frac{dx}{ax+b} = \frac{1}{a} \ln |ax+b| + C = \frac{1}{a} \ln C |ax+b|$
3. $\int \frac{x dx}{ax+b} = \frac{1}{a^2} [ax - b \ln(ax+b)] + C$
4. $\int \frac{x^2 dx}{ax+b} = \frac{x^2}{2a} - \frac{bx}{a^2} + \frac{b^2}{a^3} \ln(ax+b) + C$
5. $\int \frac{dx}{x(ax+b)} = -\frac{1}{b} \ln \left| \frac{ax+b}{x} \right| + C$
6. $\int \frac{dx}{x^2(ax+b)} = -\frac{1}{bx} + \frac{a}{b^2} \ln \left| \frac{ax+b}{x} \right| + C$
7. $\int \frac{x dx}{(ax+b)^2} = \frac{1}{a^2} \left(\ln |ax+b| + \frac{b}{ax+b} \right) + C$
8. $\int \frac{x^2 dx}{(ax+b)^2} = \frac{1}{a^3} \left[ax - 2b \ln |ax+b| - \frac{b^2}{ax+b} \right] + C$
9. $\int \frac{dx}{x(ax+b)^2} = \frac{1}{b(ax+b)} - \frac{1}{b^2} \ln \left| \frac{ax+b}{x} \right| + C$
10. $\int \frac{dx}{a^2x^2+b^2} = \frac{1}{ab} \arctan \frac{a}{b} x + C$
11. $\int \frac{dx}{a^2x^2-b^2} = \frac{1}{2ab} \ln \left| \frac{ax-b}{ax+b} \right| + C$
12. $\int \frac{x dx}{a^2x^2 \pm b^2} = \frac{1}{2a^2} \ln |a^2x^2 \pm b^2| + C$
13. $\int \frac{x^2 dx}{a^2x^2+b^2} = \frac{1}{a^2} x - \frac{b}{a^3} \arctan \frac{a}{b} x + C$
14. $\int \frac{x^2 dx}{a^2x^2-b^2} = \frac{1}{a^2} x + \frac{b}{2a^3} \ln \left| \frac{ax-b}{ax+b} \right| + C$

- $$\begin{aligned}
 15. \int \frac{dx}{x(a^2x^2+b^2)} &= \frac{1}{2b^2} \ln \frac{x^2}{a^2x^2+b^2} + C \\
 16. \int \frac{dx}{x(a^2x^2-b^2)} &= \frac{1}{2b^2} \ln \left| \frac{a^2x^2-b^2}{x^2} \right| + C \\
 17. \int \frac{dx}{x^2(a^2x^2+b^2)} &= -\frac{1}{b^2x} - \frac{a}{b^3} \arctan \frac{a}{b}x + C \\
 18. \int \frac{dx}{x^2(a^2x^2-b^2)} &= \frac{1}{b^2x} + \frac{a}{2b^3} \ln \left| \frac{ax-b}{ax+b} \right| + C
 \end{aligned}$$

II. Integrals of Irrational Functions

- $$\begin{aligned}
 19. \int (ax+b)^n dx &= \frac{(ax+b)^{n+1}}{a(n+1)} + C, \quad n \neq -1 \\
 20. \int x\sqrt{ax+b} dx &= \frac{2}{15a^2} (3ax-2b)\sqrt{(ax+b)^3} + C \\
 21. \int x^2\sqrt{ax+b} dx &= \frac{2}{105a^3} (15a^2x^2-12abx+8b^2)\sqrt{(ax+b)^3} + C \\
 22. \int \frac{x}{\sqrt{ax+b}} dx &= \frac{2(ax-2b)}{3a^2} \sqrt{ax+b} + C \\
 23. \int \frac{x^2}{\sqrt{ax+b}} dx &= \frac{2}{15a^3} (3a^2x^2-4abx+8b^2)\sqrt{ax+b} + C \\
 24. \int \frac{dx}{x\sqrt{ax+b}} &= \begin{cases} \frac{1}{\sqrt{b}} \ln \frac{\sqrt{ax+b}-\sqrt{b}}{\sqrt{ax+b}+\sqrt{b}} + C & \text{if } b > 0 \\ \frac{2}{\sqrt{-b}} \arctan \sqrt{\frac{ax+b}{-b}} + C & \text{if } b < 0 \end{cases} \\
 25. \int \frac{dx}{x^2\sqrt{ax+b}} &= -\frac{\sqrt{ax+b}}{bx} - \frac{a}{2b} \int \frac{dx}{x\sqrt{ax+b}} \\
 26. \int \frac{\sqrt{ax+b}}{x} dx &= 2\sqrt{ax+b} + b \int \frac{dx}{x\sqrt{ax+b}} \\
 27. \int \frac{\sqrt{ax+b}}{x^2} dx &= -\frac{\sqrt{ax+b}}{x} + \frac{a}{2} \int \frac{dx}{x\sqrt{ax+b}}
 \end{aligned}$$

In Formulas 28-47 it is assumed that $a > 0$, $b > 0$.

- $$\begin{aligned}
 28. \int \frac{dx}{\sqrt{a^2x^2 \pm b^2}} &= \frac{1}{a} \ln |ax + \sqrt{a^2x^2 \pm b^2}| + C \\
 29. \int \frac{dx}{\sqrt{b^2 - a^2x^2}} &= \frac{1}{a} \arcsin \frac{ax}{b} + C \\
 30. \int \frac{x dx}{\sqrt{a^2x^2 \pm b^2}} &= \frac{1}{a^2} \sqrt{a^2x^2 \pm b^2} + C
 \end{aligned}$$

31. $\int \frac{x dx}{\sqrt{b^2 - a^2 x^2}} = -\frac{1}{a^2} \sqrt{b^2 - a^2 x^2} + C$
32. $\int \frac{x^2 dx}{\sqrt{a^2 x^2 \pm b^2}} = \frac{1}{2a^3} [ax \sqrt{a^2 x^2 \pm b^2} \mp b^2 \ln |ax + \sqrt{a^2 x^2 \pm b^2}|] + C$
33. $\int \frac{x^2 dx}{\sqrt{b^2 - a^2 x^2}} = \frac{1}{2a^3} \left[-ax \sqrt{b^2 - a^2 x^2} + b^2 \arcsin \frac{ax}{b} \right] + C$
34. $\int \sqrt{a^2 x^2 \pm b^2} dx = \frac{1}{2} \left[x \sqrt{a^2 x^2 \pm b^2} \pm \frac{b^2}{a} \ln |ax + \sqrt{a^2 x^2 \pm b^2}| \right] + C$
35. $\int \sqrt{b^2 - a^2 x^2} dx = \frac{1}{2} \left[x \sqrt{b^2 - a^2 x^2} + \frac{b^2}{a} \arcsin \frac{ax}{b} \right] + C$
36. $\int x \sqrt{a^2 x^2 \pm b^2} dx = \frac{1}{3a^2} \sqrt{(a^2 x^2 \pm b^2)^3} + C$
37. $\int x \sqrt{b^2 - a^2 x^2} dx = -\frac{1}{3a^2} \sqrt{(b^2 - a^2 x^2)^3} + C$
38. $\int x^2 \sqrt{a^2 x^2 \pm b^2} dx = \frac{1}{8a^3} [ax(2a^2 x^2 \pm b^2) \sqrt{a^2 x^2 \pm b^2} - b^4 \ln |ax + \sqrt{a^2 x^2 \pm b^2}|] + C$
39. $\int x^2 \sqrt{b^2 - a^2 x^2} dx = \frac{1}{8a^3} \left[ax(2a^2 x^2 - b^2) \sqrt{b^2 - a^2 x^2} + b^4 \arcsin \frac{ax}{b} \right] + C$
40. $\int \frac{\sqrt{a^2 x^2 + b^2}}{x} dx = \sqrt{a^2 x^2 + b^2} + \frac{b}{2} \ln \frac{\sqrt{a^2 x^2 + b^2} - b}{\sqrt{a^2 x^2 + b^2} + b} + C$
41. $\int \frac{\sqrt{a^2 x^2 - b^2}}{x} dx = \sqrt{a^2 x^2 - b^2} + b \arcsin \left| \frac{b}{ax} \right| + C$
42. $\int \frac{\sqrt{b^2 - a^2 x^2}}{x} dx = \sqrt{b^2 - a^2 x^2} + b \ln \left| \frac{b - \sqrt{b^2 - a^2 x^2}}{x} \right| + C$
43. $\int \frac{\sqrt{a^2 x^2 \pm b^2}}{x^2} dx = -\frac{\sqrt{a^2 x^2 \pm b^2}}{x} + a \ln |ax + \sqrt{a^2 x^2 \pm b^2}| + C$
44. $\int \frac{\sqrt{b^2 - a^2 x^2}}{x^2} dx = -\frac{\sqrt{b^2 - a^2 x^2}}{x} - a \arcsin \frac{ax}{b} + C$
45. (a) $\int \frac{dx}{x \sqrt{a^2 x^2 + b^2}} = \frac{1}{b} \ln \left| \frac{x}{b + \sqrt{a^2 x^2 + b^2}} \right| + C$
45. (b) $\int \frac{dx}{x \sqrt{a^2 x^2 - b^2}} = -\frac{1}{b} \arcsin \left| \frac{b}{ax} \right| + C$
46. $\int \frac{dx}{x \sqrt{b^2 - a^2 x^2}} = \frac{1}{b} \ln \left| \frac{x}{b + \sqrt{b^2 - a^2 x^2}} \right| + C$
47. $\int \frac{dx}{x^2 \sqrt{a^2 x^2 \pm b^2}} = \mp \frac{\sqrt{a^2 x^2 \pm b^2}}{b^2 x} + C$

48. $\int \frac{dx}{x^2 \sqrt{b^2 - a^2 x^2}} = -\frac{\sqrt{b^2 - a^2 x^2}}{b^2 x} + C$
49. $\int \sqrt{\frac{a+x}{b+x}} dx = \sqrt{(a+x)(b+x)} + (a-b) \ln(\sqrt{a+x} + \sqrt{b+x}) + C$
50. $\int \sqrt{\frac{a-x}{b+x}} dx = \sqrt{(a-x)(b+x)} + (a+b) \arcsin \sqrt{\frac{x+b}{a+b}} + C$
51. $\int \sqrt{\frac{a+x}{b-x}} dx = -\sqrt{(a+x)(b-x)} - (a+b) \arcsin \sqrt{\frac{b-x}{a+b}} + C$
52. $\int \sqrt{\frac{1+x}{1-x}} dx = -\sqrt{1-x^2} + \arcsin x + C$
53. $\int \frac{dx}{\sqrt{(x-a)(x-b)}} = \ln \left| \frac{\sqrt{x-a} + \sqrt{x-b}}{\sqrt{x-a} - \sqrt{x-b}} \right| + C$
54. $\int \frac{dx}{\sqrt{(x-a)(b-x)}} = 2 \arcsin \sqrt{\frac{x-a}{b-a}} + C$
55. $\int x^m (ax^n + b)^{\frac{r}{s}} dx$ where m, n, r and $s (s > 1)$ are integers.
 If $\frac{m+1}{n}$ is an integer the substitution $ax^n + b = u^s$ is used and
 if $\frac{m+1}{n} + \frac{r}{s}$ is an integer the substitution $a + bx^{-n} = u^s$.
 P.L. Chebyshev proved that in all the other cases the integral
 is inexpressible in elementary functions.

III. Integrals of Transcendental Functions

56. $\int x^n e^x dx = e^x [x^n - nx^{n-1} + n(n-1)x^{n-2} - \dots + (-1)^n n!] + C$
 where n is a positive integer.
57. $\int (\ln x)^n dx = x [(\ln x)^n - n(\ln x)^{n-1} + n(n-1)(\ln x)^{n-2} - \dots + (-1)^n n!] + C$
 where n is a positive integer.
58. $\int x^n \ln x dx = \frac{x^{n+1}}{n+1} \ln x - \frac{x^{n+1}}{(n+1)^2} + C, n \neq -1$
59. $\int \frac{(\ln x)^n}{x} dx = \frac{(\ln x)^{n+1}}{n+1} + C, n \neq -1$
60. $\int \frac{dx}{x \ln x} = \ln |\ln x| + C$
61. $\int \sin mx \cos nx dx = -\frac{\cos(m+n)x}{2(m+n)} - \frac{\cos(m-n)x}{2(m-n)} + C, m \neq n$
62. $\int \sin mx \sin nx dx = \frac{\sin(m-n)x}{2(m-n)} - \frac{\sin(m+n)x}{2(m+n)} + C, m \neq n$

63. $\int \cos mx \cos nx \, dx = \frac{\sin(m-n)x}{2(m-n)} + \frac{\sin(m+n)x}{2(m+n)} + C, \quad m \neq n$
64. $\int \sin^m x \cos^n x \, dx = -\frac{\sin^{m-1} x \cos^{n+1} x}{m+n} +$
 $+ \frac{m-1}{m+n} \int \sin^{m-2} x \cos^n x \, dx = \frac{\sin^{m+1} x \cos^{n-1} x}{m+n} +$
 $+ \frac{n-1}{m+n} \int \sin^m x \cos^{n-2} x \, dx, \quad m \neq -n$
65. $\int \frac{dx}{\sin^m x \cos^n x} =$
 $= \begin{cases} -\frac{1}{m-1} \frac{1}{\sin^{m-1} x \cos^{n-1} x} + \frac{m+n-2}{m-1} \int \frac{dx}{\sin^{m-2} x \cos^{n-1} x} & \text{for } m \neq 1 \\ \frac{1}{n-1} \frac{1}{\sin^{m-1} x \cos^{n-1} x} + \frac{m+n-2}{n-1} \int \frac{dx}{\sin^m x \cos^{n-2} x} & \text{for } n \neq 1 \end{cases}$
66. $\int \tan^n x \, dx = \frac{\tan^{n-1} x}{n-1} - \int \tan^{n-2} x \, dx, \quad n \neq 1$
67. $\int x^n \sin x \, dx = -x^n \cos x + n \int x^{n-1} \cos x \, dx$
68. $\int x^n \cos x \, dx = x^n \sin x - n \int x^{n-1} \sin x \, dx$
69. $\int \frac{dx}{a+b \cos x} = \begin{cases} \frac{2}{\sqrt{a^2-b^2}} \arctan \left(\sqrt{\frac{a-b}{a+b}} \tan \frac{x}{2} \right) + C & \text{if } a^2 > b^2 \\ \frac{1}{\sqrt{b^2-a^2}} \ln \left| \frac{\sqrt{b^2-a^2} \tan \frac{x}{2} + a+b}{\sqrt{b^2-a^2} \tan \frac{x}{2} - a-b} \right| + C & \text{if } a^2 < b^2 \end{cases}$
70. $\int \frac{dx}{a+b \sin x} = \begin{cases} \frac{2}{\sqrt{a^2-b^2}} \arctan \frac{a \tan \frac{x}{2} + b}{\sqrt{a^2-b^2}} + C & \text{if } a^2 > b^2 \\ \frac{1}{\sqrt{b^2-a^2}} \ln \left| \frac{a \tan \frac{x}{2} + b - \sqrt{b^2-a^2}}{a \tan \frac{x}{2} + b + \sqrt{b^2-a^2}} \right| + C & \text{if } a^2 < b^2 \end{cases}$
71. $\int \frac{dx}{a \cos x + b \sin x} = \frac{1}{\sqrt{a^2+b^2}} \ln \left| \frac{a \tan \frac{x}{2} - b + \sqrt{a^2+b^2}}{a \tan \frac{x}{2} - b - \sqrt{a^2+b^2}} \right| + C$
72. $\int \frac{dx}{a^2 \sin^2 x + b^2 \cos^2 x} = \frac{1}{ab} \arctan \left(\frac{a \tan x}{b} \right) + C$
73. $\int e^{ax} \sin bx \, dx = \frac{e^{ax} (a \sin bx - b \cos bx)}{a^2 + b^2} + C$
74. $\int e^{ax} \cos bx \, dx = \frac{e^{ax} (b \sin bx + a \cos bx)}{a^2 + b^2} + C$

$$75. \int e^{ax} \sin^n x \, dx = \frac{e^{ax} \sin^{n-1} x (a \sin x - n \cos x)}{a^2 + n^2} + \\ + \frac{n(n-1)}{a^2 + n^2} \int e^{ax} \sin^{n-2} x \, dx$$

$$76. \int e^{ax} \cos^n x \, dx = \frac{e^{ax} \cos^{n-1} x (a \cos x + n \sin x)}{a^2 + n^2} + \\ + \frac{n(n-1)}{a^2 + n^2} \int e^{ax} \cos^{n-2} x \, dx$$

$$77. \int \arcsin x \, dx = x \arcsin x + \sqrt{1-x^2} + C$$

$$78. \int \arctan x \, dx = x \arctan x - \frac{1}{2} \ln(1+x^2) + C$$

$$79. \int x \arcsin x \, dx = \left(\frac{x^2}{2} - \frac{1}{4} \right) \arcsin x + \frac{x\sqrt{1-x^2}}{4} + C$$

$$80. \int x \arccos x \, dx = \left(\frac{x^2}{2} - \frac{1}{4} \right) \arccos x - \frac{x\sqrt{1-x^2}}{4} + C$$

$$81. \int x \arctan x \, dx = \frac{1}{2} (x^2 + 1) \arctan x - \frac{x}{2} + C$$

63. $\int \cos mx \cos nx \, dx = \frac{\sin(m-n)x}{2(m-n)} + \frac{\sin(m+n)x}{2(m+n)} + C, \quad m \neq n$
64. $\int \sin^m x \cos^n x \, dx = -\frac{\sin^{m-1} x \cos^{n+1} x}{m+n} +$
 $+\frac{m-1}{m+n} \int \sin^{m-2} x \cos^n x \, dx = \frac{\sin^{m+1} x \cos^{n-1} x}{m+n} +$
 $+\frac{n-1}{m+n} \int \sin^m x \cos^{n-2} x \, dx, \quad m \neq -n$
65. $\int \frac{dx}{\sin^m x \cos^n x} =$
 $= \begin{cases} -\frac{1}{m-1} \frac{1}{\sin^{m-1} x \cos^{n-1} x} + \frac{m+n-2}{m-1} \int \frac{dx}{\sin^{m-2} x \cos^n x} & \text{for } m \neq 1 \\ \frac{1}{n-1} \frac{1}{\sin^{m-1} x \cos^{n-1} x} + \frac{m+n-2}{n-1} \int \frac{dx}{\sin^m x \cos^{n-2} x} & \text{for } n \neq 1 \end{cases}$
66. $\int \tan^n x \, dx = \frac{\tan^{n-1} x}{n-1} - \int \tan^{n-2} x \, dx, \quad n \neq 1$
67. $\int x^n \sin x \, dx = -x^n \cos x + n \int x^{n-1} \cos x \, dx$
68. $\int x^n \cos x \, dx = x^n \sin x - n \int x^{n-1} \sin x \, dx$
69. $\int \frac{dx}{a+b \cos x} = \begin{cases} \frac{2}{\sqrt{a^2-b^2}} \arctan \left(\sqrt{\frac{a-b}{a+b}} \tan \frac{x}{2} \right) + C & \text{if } a^2 > b^2 \\ \frac{1}{\sqrt{b^2-a^2}} \ln \left| \frac{\sqrt{b^2-a^2} \tan \frac{x}{2} + a+b}{\sqrt{b^2-a^2} \tan \frac{x}{2} - a-b} \right| + C & \text{if } a^2 < b^2 \end{cases}$
70. $\int \frac{dx}{a+b \sin x} = \begin{cases} \frac{2}{\sqrt{a^2-b^2}} \arctan \frac{a \tan \frac{x}{2} + b}{\sqrt{a^2-b^2}} + C & \text{if } a^2 > b^2 \\ \frac{1}{\sqrt{b^2-a^2}} \ln \left| \frac{a \tan \frac{x}{2} + b - \sqrt{b^2-a^2}}{a \tan \frac{x}{2} + b + \sqrt{b^2-a^2}} \right| + C & \text{if } a^2 < b^2 \end{cases}$
71. $\int \frac{dx}{a \cos x + b \sin x} = \frac{1}{\sqrt{a^2+b^2}} \ln \left| \frac{a \tan \frac{x}{2} - b + \sqrt{a^2+b^2}}{a \tan \frac{x}{2} - b - \sqrt{a^2+b^2}} \right| + C$
72. $\int \frac{dx}{a^2 \sin^2 x + b^2 \cos^2 x} = \frac{1}{ab} \arctan \left(\frac{a \tan x}{b} \right) + C$
73. $\int e^{ax} \sin bx \, dx = \frac{e^{ax} (a \sin bx - b \cos bx)}{a^2 + b^2} + C$
74. $\int e^{ax} \cos bx \, dx = \frac{e^{ax} (b \sin bx + a \cos bx)}{a^2 + b^2} + C$

-
18. Piskunov, N. S., *Differential and Integral Calculus*. Mir Publishers, Moscow, 1974.
 - 19*. Romanovsky, P. I., *Fourier's Series. Field Theory. Analytic and Special Functions. Laplace's Transformation*. "Nauka", Moscow, 1973.
 - 20*. Smirnov, V. I., *A Course in Higher Mathematics*, Vols. 1, 2. "Nauka", Moscow, 1965.
 21. Tikhonov, A. N., Samarsky, A. A., *Equations of Mathematical Physics*. "Nauka", Moscow, 1966 (in Russian). The English translation: *Equations of Mathematical Physics*. New York, Pergamon Press, 1963.
 - 22*. Tolstov, G. P., *Elements of Mathematical Analysis*, Vols. 1, 2. "Nauka", Moscow, 1966.
 23. Yefimov, N. V., *A Brief Course in Analytic Geometry*. Mir Publishers, 1964.
 - 24*. Yefimov, N. V., *Quadratic Forms and Matrices*. "Nauka", Moscow, 1972.

NAME INDEX

Abel, N. H. 248, 685, 717
Akiŭis, M. A. 767
Ampère, A. M. 557
Angot, M. A. 767
Aramanovich, I. G. 767

Beklemishev, D. A. 767
Bermant, A. F. 5
Bernoulli, Jacob 579
Bernoulli, Johann 579
Bessel, F. W. 751
Bézout, E. 245
Biot, J. B. 557
Bronstein, I. N. 767

Cardano, G. 248
Cauchy, A. 180, 566, 594, 678, 759
Chebyshev, P. L. 749, 765
Clairaut, A. 592
Clapeyron, B. P. E. 380, 510
Coulomb, C. A. 508

D'Alembert, J. 668
Demidovich, B. P. 767
De Moivre, A. 239
Descartes, R. 16
Dirichlet, P. G. L. 340, 731

Eisgolts, L. E. 767
Euclid 15
Euler, L. 97, 242, 327, 456, 478, 480

Fermat, P. 175
Ferrari, L. 248
Fikhtengolts, G. M. 767
Fourier, J. B. J. 724

Galilei, G. 573
Galois, E. 248

Gauss, K. F. 531, 556
Goldberg, V. V. 767
Golovina, L. I. 767
Green, G. 491
Guter, R. S. 767

Hamilton, W. R. 547
Hamming, R. V. 767
Höené-Wronski, J. M. 623, 655
Hooke, R. 35, 628

Jacobi, K. G. J. 451

Korn, G. A. 767
Korn, T. M. 767
Kudryavtsev, L. D. 767

Lagrange, J. L. 178, 423, 619, 696
Laplace, P. S. 554, 557, 579, 712
Leibniz, G. 16, 126, 171, 314, 396, 474, 475, 502, 674
Leverrière, J. 18
Levin, V. I. 767
L'Hospital, J. F. A. 204
Lobachevsky, N. I. 255
Lunts, G. L. 767
Lyapunov, A. M. 751

Maclaurin, C. 699
Maron, I. A. 767
Maxwell, J. C. 561
Mendelev, D. I. 380, 510
Möbius, A. F. 521
Mohammed Al-Khorezmi 15
Myškis, A. D. 767

Napier, J. 60
Nesterov, P. N. 18
Newton, I. 16, 123, 314, 502

Ostrogradsky, M. V. 521
Ovchinsky, B. V. 767

Parseval, M. A. 751
Piskunov, N. P. 767
Poisson, S. D. 327, 456, 557
Pythagoras, 15

Riccati, J. 582
Rolle, M. 176
Romanovsky, P. I. 767

Samarsky, A. A. 767
Savart, F. 557
Semendyaev, K. A. 767
Shuvalova, E. Z. 767
Simpson, T. 323

Smirnov, V. I. 767
Stirling, J. 115
Stokes, G. G. 528

Taylor, B. 691
Tikhonov, A. N. 767
Tolstov, G. P. 767
Torricelli, E. 573
Tslaf, L. Ya. 6
Tumarkin, L. A. 6

Vieta, F. 15

Yefimov, N. V. 767

Zhukovsky, N. E. 18

SUBJECT INDEX

- Abel's theorem 685, 717
- Absolute
 - error 24
 - value 22, 238
- Absolutely convergent
 - integral 339
 - series 677
- Absolutely integrable (summable) function 754
- Acceleration 236
 - centripetal (normal) 237
 - of convergence of a series 674
 - of rectilinear motion 168
 - tangential 237
- Acyclicity 505, 552
- Additivity of the integral 304, 359
- Adiabatic process 512
- Algebraic
 - equation 245
 - multiple root of 246
 - not solvable by radicals 248
 - simple root of 246
 - function 35f, 38
- Alternating series 674
- Ampère's rule (right-hand screw rule) 557f
- Amplitude 62, 629
- Analytic geometry 16
- Analytical
 - function 717
 - representation of functions 29
- Anchor ring (torus) 350
- Angle
 - between intersecting curves 148
 - of contingency 221
 - solid 556
- Antiderivative (primitive) 258
- Aperiodic motion 628
- Approximate
 - calculations 23ff
 - integration 323ff
 - methods for solving differential equations 581ff
 - solution of equations 249f
 - value of a magnitude 24
- Arc length 219, 350ff
 - differential of 220
- Area
 - in Cartesian coordinates 345f
 - of a curvilinear trapezoid 291ff
 - in polar coordinates 346ff
 - of a surface of revolution 362f
- Argument 26
 - of a complex number 238
- Arithmetic mean of a function 310f
- Asymptote 211ff
- Autonomous system
 - of differential equations 640
- Average
 - curvature 221
 - velocity 118
- Basic
 - elementary functions 32
 - table of integrals 262
- Bernoulli's differential equation 579
- Bessel's inequality 751
- Beta function (Euler's integral of the first kind) 480
- Bézout's theorem 245
- Binomial series 702
- Blot-Savart-Laplace law 557
- Boundary conditions 594
- Boundary-value problem 594
- Bounded function 82
- Branch of a multiple-valued function 37, 54f
- Cardano's solution of the cubic 247, 248
- Catenary 598
- Cauchy's
 - existence and uniqueness theorem for solutions of differential equations 566, 594, 602, 639
 - integral test 671
 - principal value of a divergent integral 759
 - problem 566, 594, 602

- product of series 678
- root test 670
- theorem 180
- Centre of gravity 355, 458f, 459
- Characteristic equation 607, 625, 645, 656
- Characteristic root 607
- Chebyshev's polynomials 749
- Circular frequency 63
- Circular helix (screw line) 228
- Clairaut's differential equation 592
- Closed contour of integration, orientation of 490
- Commensurable and incommensurable frequencies 725
- Common (decimal) logarithms 60
- Commutativity property of an absolutely convergent series 677
- Comparison
 - of infinitesimals 109ff
 - of infinitely large magnitudes 114ff
 - tests for convergence of a series 666
- Complete elliptic integrals of the 1st and 2nd kinds 353
- Complex
 - amplitudes, method of 618f
 - exponential function 242, 718
 - function of a real argument 241
 - number 237ff
 - plane 238
- Composite function 33
- Conditionally convergent series 677
- Conjugate complex number 239
- Constant of variation 44
- Constraint 421
- Continuity 98ff, 232
- Contour lines 369
- Convection 561
- Convective term (convective derivative) 561
- Convergence in the mean 752
- Coordinate
 - line 442
 - surface 465
- Coordinates
 - curvilinear 451
 - cylindrical 465
 - spherical 466f
- Coulomb's law 508
- Coupling equation 421
- Cubical parabola 52
- Curvature 222
 - average 221
 - centre of 224
 - circle of 224
 - radius of 224, 236
 - of a space curve 236
- Curve
 - adiabatic 512
 - concave 199
 - convex 198
 - osculating plane of 237
 - piecewise smooth 487, 731
 - principal normal of 236
 - rectifiable 352
 - simple 371, 490
 - smooth 220, 227, 351, 487
- Curvilinear
 - coordinates 235, 451
 - sector 346
 - trapezoid 291
- Cusp locus 591
- Cut-and-try method 250f
- Cycle 510
- Cycloid 146
- Cylindroid 437
- D'Alembert's test 668f
- Damped oscillations 628f
- Decomposition of a rational fraction into partial fractions 276ff
- Definite integral 298
 - differentiation with respect to the upper limit of integration of 312ff
 - estimation of 306f, 308f
 - general scheme for application of 358ff
 - geometrical meaning of 305f
 - integration by change of variable (by substitution) in 319f
 - integration by parts in 318f
 - interchanging limits in 302f
 - sign of 304
 - simplest properties of 300ff
 - splitting interval of integration of 303f
 - theorem on existence of 300
- De Moivre's formula 239
- Density 120, 457
- Dependent variable 26
- Derivative(s) 123
 - of basic elementary functions 134ff
 - of complex function 241
 - of composite function 131, 400f
 - convective 561
 - of elementary functions 139ff
 - geometrical meaning in Cartesian coordinates of 125
 - geometrical meaning in polar coordinates of 153
 - of higher orders 167
 - of implicit function 141f
 - of inverse function 133
 - left-hand 162

- logarithmic 141
- one-sided 162f
- of power function 124
- partial 376, 390f
 - explicit 402
 - geometrical meaning of 378
 - of higher order 390f
 - mixed 391
 - total 402
- right-hand 163
- of vector function 333
- Diameter of a domain 438
- Differentiable function 160, 383f
- Differential 156
 - application to approximate calculations of 163f
 - of arc length 220
 - of composite function 158
 - of higher order 392f
 - invariance of 158
 - of the n th order 171, 392f
 - partial 379, 392f
 - properties of 157f
 - total 382, 386f
- Differential equation(s) 564
 - approximate integration of 583f
 - with constant coefficients 607f, 611f, 644f
 - direction field of 582
 - exact 581
 - of the first order 564
 - fundamental system of solution of 624, 643, 655
 - general integral of 568, 577
 - general solution of 566
 - of higher orders 602f
 - nonhomogeneous 575
 - homogeneous linear 603, 622
 - integral curve of 567
 - integrable by quadratures 569
 - isocline of 582
 - linear 576, 603f, 622f, 625f
 - nonhomogeneous linear 603, 622
 - numerical integration of 524f
 - order of 602
 - ordinary 565
 - partial 565
 - particular solution of 566
 - of the second order 592
 - singular point of 585, 604
 - singular solution of 588
 - system of 635f
 - in normal form 636
 - with variable separable 568f
- Differentiation 126
 - of basic elementary functions 134f
 - basic formulas of 139
 - of composite function 131, 400f
 - connection with arithmetical operations of 127f
 - of elementary functions 133f
 - graphical 151f
 - of implicit functions 141f, 407f
 - of inverse function 133
 - parametric 147, 169f
 - rules 139
- Direct variation 44
- Directional derivative 429
- Dirichlet's
 - condition 731
 - integral 340
- Discontinuity 101
 - of the first kind 101f
 - infinite 101
 - removable 102
 - of the second kind 103
- Domain 370, 375
 - boundary of 371
 - bounded 371
 - closed 371
 - connected 371
 - of definition of a function 25, 366
 - of integration 439
 - multiply connected 372
 - open 370
 - simply connected 371, 565
 - starlike (star-shaped) 454, 555
 - unbounded 371
- Dominant series 680
- Dominated series 680
- Double integral 439
 - applications of 457f
 - change of variables in 450f
 - existence theorem for 440
 - in polar coordinates 452f
 - properties of 440f
 - reduction to an iterated integral of 445
- Eigenvalue 655
- Eigenvector 655
- Element of area
 - in Cartesian coordinates 442
 - in curvilinear coordinates 451
 - in polar coordinates 453
- Element of integration 251, 296, 439
- Element volume
 - in Cartesian coordinates 452
 - in cylindrical coordinates 455
 - in spherical coordinates 456
- Entropy 511
- Envelope 529f
- Equation of state (in thermodynamics) 509

- Equipotential surface 427
 Equivalent
 (asymptotically) infinitely large magnitudes 114
 infinitesimals 110
 Error 24f
 Euler-Poisson integral 327, 456
 Euler's
 formula 242, 244, 718
 integral of the first kind (beta function) 480
 integral of the second kind (gamma function) 478
 method for approximate solution of differential equations 583f
 Even function 38
 Evolvent (Involute) 224f
 Evolute 224
 Exciting (disturbing) force 627
 Existence and uniqueness theorem for solutions of differential equations 566, 594, 639
 Exponential form of a complex number 243
 Exponential function 57, 242, 718
 Extremum 185, 415
 absolute 185
 conditional 420
 necessary conditions for 186, 415
 non-strict 45, 185
 relative 185
 strict 185
 sufficient conditions for 187, 194, 418
 unconditional 420
 Fermat's theorem 175
 Ferrari's solution of the quartic 248
 Field
 electromagnetic 554ff
 electrostatic 508f
 of force 482, 504
 intensity of 482
 potential 507
 potential function of 507
 work of 482ff, 507
 of gravitation 507f
 of linear velocities 534
 magnetic, generated by an electric current 537ff
 nonstationary 559ff
 plane 428, 431
 scalar 427
 directional derivative of 429
 gradient of 431
 level line of 428
 level surface of 427
 stationary 427, 480, 504, 533
 vector 427, 480, 516, 533ff
 circulation of 543
 divergence of 539ff
 flux across a surface of 537
 harmonic 553
 homogeneous 534
 irrotational 551
 plane 534, 536
 potential 550
 potential function of 550
 rotation (curl) of 546
 sink of 538f
 solenoidal 551
 source of 538f
 stationary 533
 tube of 552
 vector line of 535
 vector-potential of 553
 Fluid
 flow
 plane 534
 stationary 515ff
 streamline of 535
 flux across a surface of 516ff, 537
 pressure on the wall of the vessel 363f
 Focal point of a differential equation 588
 Force
 disturbing (exciting) 627
 elastic 628
 resisting 627
 restoring 627
 Forced oscillations 628, 630ff
 Fourier's
 coefficients 730, 748
 generalized 748
 minimizing property of 751
 with respect to an orthonormal system of functions 748
 cosine series 734
 cosine transform 760
 integral 765
 in complex form 258ff
 for even and odd functions 756ff
 integral formula 765
 inverse transform 759
 polynomial 732
 series 730, 748
 in complex form 740ff
 of an even function 732
 expansion theorem for a function into 731
 of a function with arbitrary period 733f
 generalized 748
 half-range (incomplete) 734

- of an odd function 733
 - with respect to an orthonormal system of functions
- sine series 734
- sine transform 760
- transform 759
- transformation 760
- Free (natural) oscillations 628
- Frequency 62
- Function(s) 26
 - absolutely integrable (summable) 754
 - algebraic 35f, 38
 - analytical 717
 - basic elementary 32
 - bounded 82f
 - composite 33
 - continuous 98
 - of a continuous argument 27
 - decreasing 40
 - derivative of 123
 - differentiable 160, 382
 - domain of definition of 26
 - elementary 33f
 - entire rational 35
 - even 38
 - explicit 36
 - exponential 57
 - greatest value of 41
 - harmonic 554
 - hyperbolic 67
 - implicit 36
 - increasing 40
 - infinite-valued 65
 - of an integral argument 27
 - inverse 51
 - inverse hyperbolic 69f
 - inverse trigonometric 64
 - least value of 41
 - limit of 72, 75
 - linear 44
 - linear-fractional 49
 - logarithmic 58
 - methods of representing of 28f, 366f
 - monotone 41
 - multiple-valued 37, 51
 - nonelementary 35
 - odd 39
 - periodic 39
 - piecewise smooth 730
 - point of discontinuity of 100
 - power 55
 - quadratic 27, 46
 - rate of change of 123
 - of several variables 375f
 - conditional extremum of 420f
 - reconstruction from the total differential of 397f
 - smooth 730
 - stationary point of 188
 - transcendental 36
 - trigonometric (circular) 60
 - of two variables 365f
 - continuity of 373
 - determining the greatest and the least values of 419f
 - extremum of 415f
 - implicit 367
 - level line of 369
 - limit of 372
 - method of representing of 366f
 - partial derivative of 376
 - points of discontinuity of 374
 - reconstruction from the total differential of 393f
 - vector 230
- Functional 300
- Functional determinant (Jacobian) 451
- Functional relationship 17
- Functional series 678
- Fundamental system of solutions 642, 643, 655
- Gamma function (Euler's integral of the second kind) 478
- Gauss-Ostrogradsky theorem (formula) 531
 - in vector form 540
- Gauss' theorem (in electrostatics) 556
- General scheme for investigating the graph of a function 216f
- Geodesic 230
- Geometrical series 661
- Graph
 - of a function 30, 367
 - of motion 118
- Graphical
 - differentiation 151f
 - integration 329f
 - representation of functions 30, 367
- Greatest value of a function 41, 195, 419
- Green's formula (theorem) 419f
- Half-life of a radioactive substance 572
- Half-plane 371
- Half-space 505
- Hamiltonian operator (or nabla or del) 547
 - repeated operations with 549f
- Harmonic
 - analysis 63

- compound 64
- functions 554
- simple 63, 724
- vibration 62
- Heat capacity 121, 510f
- Hodograph 231, 241
- Hooke's law 35, 628
- Horizontal 370
- Hyperbola of order β 56
- Hyperbolic functions 67, 719
- Hyperbolic substitution 274f
- Ideal (perfect) gas 510
- Imaginary axis 238
- Implicit function 36, 367, 405
 - theorem on the existence of 405, 407
- Improper integral 332
 - dependent on a parameter 476
 - regularly convergent 476
- Infinite 332
 - absolutely convergent 339
 - comparison test for 336
 - conditionally convergent 340
 - convergent 332
 - divergent 332
 - of a function with infinite discontinuity 340ff
 - convergent 341
 - divergent 341
 - tests for convergence of 335ff
- Inclusion relation 22
- Increment 45, 379, 381
- Indefinite integral(s) 261
 - basic table of 262
 - integration by change of variable (by substitution) in 272f
 - integration by parts in 269
- Indefinite integration 261
 - in elementary functions 290
 - of irrational functions 282ff
 - of rational functions 275f
 - simplest rules for 263ff
 - of trigonometric functions 284f
- Independent variable 26
- Indeterminate forms 209
- Infinite integral 332
- Infinitely large magnitude 80
- Infinitesimal 84
- Infinite-valued function 65
- Initial condition(s) 566, 594, 602
- Initial phase of a harmonic vibration 62, 629
- Integral
 - definite 298
 - dependent on a parameter 471f
 - improper 476ff
 - double 439
 - improper 332
 - indefinite 261
 - line 484ff
 - surface 518
 - triple 549
- Integral curve 261, 331, 567
- Integral sum 297, 439, 460, 484, 512, 518
 - limit of 297
- Integrand 261, 298, 439
- Integrating 331
- Interior point 370
- Intermediate variable 33
- Interval 21
 - of integration 298
- Invariance of the form of the first differential 158, 403ff
- Inverse
 - function 51
 - hyperbolic functions 69f
 - trigonometric functions 64
 - variation 48
- Irrational equation 248
- Irrational numbers 20
- Isobar 370
- Isocline 582
- Isotherm 370
- Iterated (repeated) integral 445
- Iteration method 255
- Jacobian (functional determinant) 451
- Lagrange's
 - form of the remainder of a power series 696
 - method of variation of arbitrary constants 619ff
 - multipliers 423, 425
 - theorem of finite increments 179
- Laplace's
 - equation 554
 - function (probability integral) 579, 712
- Laplacian operator 549f
- Least value of a function 41, 195, 419
- Leibniz'
 - rule for differentiating a product 171
 - rule for differentiating an integral with respect to a parameter 396, 474, 475
 - test for convergence of an alternating series 674
- Level line 369
- Level surface 427
- L'Hospital's rule 204

- Limit(s)
 of a function 72, 75, 231
 of the function $\frac{\sin x}{x}$ as $x \rightarrow 0$ 94
 general rules for finding 85f
 of integration 298
 one-sided 101
 of a sequence 78
 of the sequence
 $\left(1 + \frac{1}{n}\right)^n$ as $n \rightarrow \infty$ 96
 Limiting error 24f
 Linear
 combination 45, 604, 622
 differential equation 576, 603ff, 622ff, 625ff
 function 44
 Linear-fractional function 49
 Linearly dependent and independent systems of functions 622, 654
 Line integral 484
 applications of 507ff
 over a closed curve 490f
 evaluation of 485ff, 513
 of the first type 484, 512
 path-independence of 494ff, 505
 with respect to arc length 484, 512f
 with respect to coordinates 484, 504
 of the second type 484, 504
 over a space curve 503ff
 Line of force 535
 Lobachevsky's method 255
 Logarithmic
 curve 58
 decrement 630
 differentiation 140ff
 function 58
 Lyapunov's theorem 752

 Maclaurin's series 700
 expanding functions into 699ff
 Magnitude 16
 Mapping 51, 452
 many-to-one 51
 one-to-one 51
 Mass
 of a material line 295f
 of a nonhomogeneous body 459f
 of a plate 457f
 Mathematical analysis 15
 Matrix form of a linear system of differential equations 652ff
 coefficient matrix of 653
 Maximum of a function 181, 194
 Maxwell's equations 561

 Mean
 square deviation 750
 value of a function 310, 442
 value theorem 309, 442
 Mendeleev-Clapeyron equation 380, 510
 Method
 of chords 251ff
 combined 254f
 of complex amplitudes 618f
 cut-and-try 250
 Euler's 583f
 iteration 255
 of Lagrange's multipliers 423, 425
 of least squares 750
 Lobachevsky's 255
 operational 635
 of parallel sections 368f
 of tangents 253f
 of undetermined coefficients 612ff
 Minimum of a function 185, 194
 Möbius strip 521
 Modulus of a complex number 238
 Moment of inertia 459, 469
 Monotonicity 41, 181ff

 Nabla (or del or Hamiltonian operator) 547
 Natural
 (free) vibration 628
 frequency 627, 629
 n-dimensional space 376, 642
 (Napierian) logarithms 60, 113
 numbers 19
 Neighbourhood of a point 21, 370, 375f
 Newton-Leibniz theorem (formula) 314, 502, 506
 Nodal point of a differential equation 587
 Non-elementary function 35
 Normal
 to a curve 148, 236
 to a surface 410
 Normal plane to a curve 226
 Number e 98
 Number scale 19
 Numerical methods of integration of differential equations 484f

 Odd function 39
 One-parameter family of curves, envelope of 589ff
 One-to-one
 correspondence 20, 51
 mapping 51
 Operational calculus 627

- Operations on continuous functions 103f
 Orders
 of infinitesimals 109, 112
 of infinitely large magnitudes 114
 Orthogonal system of functions 746
 Orthonormal system of functions 747
 Osculating plane 237
- Parabola of order α 56
 Parametric
 equations
 of a curve 143
 of a cycloid 147
 of an ellipse 142
 representation of a function 143
 Parseval's relation 752
 Partial derivative 376
 Partial differential 379
 Particular solution of a differential equation 566
 Partition of a domain 438
 finess of 439
 Passing to limit in Inequalities 92f
 Path of integration 485
 Path length, determination of 295f
 Period
 of a function 39
 primitive 40, 725
 of oscillations 62
 Periodic function 39
 Phase
 plane 639
 space 642
 spectrum 742, 760
 trajectory 640
 of vibrations 62, 629
 Plane layer 505
 Planimeter 331
 Point-by-point plotting of graphs 42
 Point
 of conditional extremum 420
 of discontinuity 101f, 347
 of inflection 200
 of maximum 184, 415
 of minimum 185, 415
 Point function 427
 Poisson's equation 557
 Polynomial 35
 Polytropic curves 57
 Potential of the force of attraction 35
 Power function 55
 Power series 684
 Practical harmonic analysis (or wave analysis or Fourier analysis) 742f
 Primitive 258, 502, 506
 Principal normal 236
- Products
 of inertia 459, 470
 Properties of continuous functions 107f
p series
- Quadratic function 27, 46
 Qubical parabola 52
- Radioactive decay 571f
 Radius of convergence 687, 717
 Radius of curvature 224
 Rate of change of a function 121f
 Rational
 equation 248
 fraction
 partial 276
 proper 276
 numbers 19
 Reaction rate 121
 Real axis 238
 Real numbers 19
 Rectangle rule 324
 Rectification of a curve 352
 Recurrence formulas 281, 287, 319, 659
 Regularly convergent
 functional series 680
 improper integral 476
 Relative error 25
 Removable singularity 102
 Resistance coefficient 627
 Resisting force 627
 Resonance 631f
 curve 633
 Restoring force 627
 Riccati's differential equation 582
 Right-hand screw rule (Ampère's rule) 558
 Rolle's theorem 176
 Root (zero) of a function 38
 Rotation (curl) of a vector (field) 546
- Saddle point of a differential equation 587
 Screw line 228
 Secular term 634
 Semicubical parabola 56
 Sequence 78
 bounded 95
 Series
 Fourier's 730
 functional 678
 continuity of the sum of 682
 domain of convergence of 678
 point of convergence of 678

- Limit(s)
 of a function 72, 75, 231
 of the function $\frac{\sin x}{x}$ as $x \rightarrow 0$ 94
 general rules for finding 85ff
 of integration 298
 one-sided 101
 of a sequence 78
 of the sequence
 $\left(1 + \frac{1}{n}\right)^n$ as $n \rightarrow \infty$ 96
 Limiting error 24f
 Linear
 combination 45, 604, 622
 differential equation 576, 603ff,
 622ff, 625ff
 function 44
 Linear-fractional function 49
 Linearly dependent and independent
 systems of functions 622, 654
 Line integral 484
 applications of 507ff
 over a closed curve 490f
 evaluation of 485ff, 513
 of the first type 484, 512
 path-independence of 494ff, 505
 with respect to arc length 484, 512f
 with respect to coordinates 484, 504
 of the second type 484, 504
 over a space curve 503ff
 Line of force 535
 Lobachevsky's method 235
 Logarithmic
 curve 58
 decrement 630
 differentiation 140ff
 function 58
 Lyapunov's theorem 752

 Maclaurin's series 700
 expanding functions into 699ff
 Magnitude 16
 Mapping 51, 452
 many-to-one 51
 one-to-one 51
 Mass
 of a material line 295f
 of a nonhomogeneous body 459f
 of a plate 457f
 Mathematical analysis 15
 Matrix form of a linear system of diffe-
 rential equations 652ff
 coefficient matrix of 653
 Maximum of a function 181, 194
 Maxwell's equations 561

 Mean
 square deviation 750
 value of a function 310, 442
 value theorem 309, 442
 Mendeleev-Clapeyron equation 380, 510
 Method
 of chords 251ff
 combined 254f
 of complex amplitudes 618f
 cut-and-try 250
 Euler's 583f
 iteration 255
 of Lagrange's multipliers 423, 425
 of least squares 750
 Lobachevsky's 255
 operational 635
 of parallel sections 368f
 of tangents 253f
 of undetermined coefficients 612ff
 Minimum of a function 185, 194
 Möbius strip 521
 Modulus of a complex number 238
 Moment of inertia 459, 469
 Monotonicity 41, 181ff

 Nabla (or del or Hamiltonian operator)
 547
 Natural
 (free) vibration 628
 frequency 627, 629
 n-dimensional space 376, 642
 (Napierian) logarithms 60, 113
 numbers 19
 Neighbourhood of a point 21, 370, 375f
 Newton-Leibniz theorem (formula)
 314, 502, 506
 Nodal point of a differential equation
 587
 Non-elementary function 35
 Normal
 to a curve 148, 236
 to a surface 410
 Normal plane to a curve 226
 Number e 98
 Number scale 19
 Numerical methods of integration of
 differential equations 484f

 Odd function 39
 One-parameter family of curves, en-
 velope of 589ff
 One-to-one
 correspondence 20, 51
 mapping 51
 Operational calculus 627

- Operations on continuous functions 103ff
 Orders
 of infinitesimals 109, 112
 of infinitely large magnitudes 114
 Orthogonal system of functions 746
 Orthonormal system of functions 747
 Osculating plane 237
- Parabola of order α 56
 Parametric
 equations
 of a curve 143
 of a cycloid 147
 of an ellipse 142
 representation of a function 143
 Parseval's relation 752
 Partial derivative 376
 Partial differential 379
 Particular solution of a differential equation 566
 Partition of a domain 438
 finess of 439
 Passing to limit in inequalities 92f
 Path of integration 485
 Path length, determination of 295f
 Period
 of a function 39
 primitive 40, 725
 of oscillations 62
 Periodic function 39
 Phase
 plane 639
 space 642
 spectrum 742, 760
 trajectory 640
 of vibrations 62, 629
 Plane layer 505
 Planimeter 331
 Point-by-point plotting of graphs 42
 Point
 of conditional extremum 420
 of discontinuity 101f, 347
 of inflection 200
 of maximum 184, 415
 of minimum 185, 415
 Point function 427
 Poisson's equation 557
 Polynomial 35
 Polytropic curves 57
 Potential of the force of attraction 35
 Power function 55
 Power series 684
 Practical harmonic analysis (or wave analysis or Fourier analysis) 742ff
 Primitive 258, 502, 506
 Principal normal 236
- Products
 of inertia 459, 470
 Properties of continuous functions 107ff
 p series
- Quadratic function 27, 46
 Cubical parabola 52
- Radioactive decay 571f
 Radius of convergence 687, 717
 Radius of curvature 224
 Rate of change of a function 121f
 Rational
 equation 248
 fraction
 partial 276
 proper 276
 numbers 19
 Reaction rate 121
 Real axis 238
 Real numbers 19
 Rectangle rule 324
 Rectification of a curve 352
 Recurrence formulas 281, 287, 319, 659
 Regularly convergent
 functional series 680
 improper integral 476
 Relative error 25
 Removable singularity 102
 Resistance coefficient 627
 Resisting force 627
 Resonance 631ff
 curve 633
 Restoring force 627
 Riccati's differential equation 582
 Right-hand screw rule (Ampère's rule) 558
 Rolle's theorem 176
 Root (zero) of a function 38
 Rotation (curl) of a vector (field) 546
- Saddle point of a differential equation 587
 Screw line 228
 Secular term 634
 Semicubical parabola 56
 Sequence 78
 bounded 95
 Series
 Fourier's 730
 functional 678
 continuity of the sum of 682
 domain of convergence of 678
 point of convergence of 678

- permanently convergent 686
- regularly convergent 680
- termwise integration and differentiation of 682f
- uniformly convergent 681
- general term of 659
- numerical 659
 - absolutely convergent 677, 716
 - alternating 674
 - with arbitrary terms, test for convergence of 676ff
 - Cauchy's integral test for convergence of 671
 - Cauchy's root test for convergence of 670
 - with complex terms 716
 - conditionally convergent 677
 - convergent 660
 - D'Alembert's test for convergence of 668
 - divergent 660
 - general term of 659
 - harmonic 664
 - necessary condition for convergence of 664
- partial sum of 660
- positive 665
 - divergent to $+\infty$ 666
 - test for convergence of 665ff, 670ff
- power 684
 - binomial 702
 - circle of convergence of 717
 - coefficients of 684
 - in complex argument 716ff
 - expansion of functions into 691ff, 694f, 699ff
 - interval of convergence of 687
 - Maclaurin's 700
 - properties of 689ff
 - radius of convergence of 687, 717
 - Taylor's 693
 - in two variables 719ff
- remainder of 662
- sum of 660
- trigonometric 726
- Signum function 74
- Simple harmonic 63, 724
- Simpson's rule 327
- Sine integral 712
- Singular point
 - of a curve 226
 - of a differential equation 585, 641
 - of a surface 410
- Spectral
 - function 760
 - sequence 741, 760
- Spectrum
 - amplitude 742, 760
 - continuous 756
 - discrete 755
 - phase 742, 760
- Spring constant 628
- Stability theory 641
- Static moment 355, 458
- Stationary point 188, 416
- Stiffness factor 628
- Stirling's formula 115
- Stokes
 - law 599
 - theorem (formula) 505, 528, 546
- Stream line (flow line) 535
- Strip 371
- Sufficient condition for differentiability of a function 384
- Superposition of simple harmonics 64, 724
- Surface
 - normal to 410
 - one-sided 521
 - smooth 522
 - two-sided 520
- Surface integral 518
 - over a closed surface 521f
 - evaluation of 522ff
 - of the first type 526
 - properties of 519ff
 - of the second type 526
- System of differential equations 635ff
 - autonomous 640
 - geometrical, mechanical and hydrodynamical interpretation of solutions of 639ff, 642
 - linear 642ff
 - rest point of 641
 - singular point of 641
 - for vector lines 536
- Tabular representation of functions 28
- Tangent plane to a surface 386
- Tangent to a curve 125, 148, 225
- Taylor's
 - coefficient 693
 - formula 694f
 - applications of 706ff
 - Lagrange's form of the remainder of 696, 698
 - remainder of 694, 720
 - polynomial 694
 - series 693, 720
 - expanding functions into 699ff
- Terminal velocity of fall 600
- Test(s)
 - for convergence of a series 676

- for convexity and concavity of a curve 200f
- for determination of points of inflection 201f
- for existence of the limit of sequence 95
- for extremum 186ff
- for monotonicity 181f
- Theory of probability 115, 219
- Thermal capacity 510f
- Thermodynamical diagram 510
- Thermodynamics, problems of 509ff
- Three-leafed rose, area of 347f
- Time derivative
 - explicit partial 560
 - total (particle) 561
- Toricelli's law 573
- Torus (anchor ring) 350
- Total differential 382
 - application of 388ff
 - geometrical interpretation of 386ff
 - integration of 501f
- Trajectory
 - of motion 118
 - of a system of differential equations 639
- Transcendental equation 248
- Transcendental functions 36
- Transient process 634
- Trapezoid formula 324
- Trigonometric
 - form of a complex number 239
 - functions 60, 719
 - series 726
 - substitution 274f
- Triple integral 459ff
 - applications of 468ff
 - in Cartesian coordinates 461ff
 - in cylindrical coordinates 465f
 - evaluation of 461f
 - in spherical coordinates 466
- Twelve-ordinate scheme 744
- Uniform deviation of functions 749
- Uniformly convergent functional series 681
- Variable 16
 - dependent 26
 - independent 26
 - of integration 261, 439
- Vector
 - field 516, 533ff
 - function 230, 653
 - coordinates of 653
- Velocity 118ff, 235
- Vibration(s) 627ff
 - amplitude of 62, 629
 - damped harmonic 629
 - in electric circuits 634f, 648f
 - forced 628, 631
 - free (natural) 628, 634
 - harmonic 62, 629
 - initial phase of 62, 629
 - logarithmic decrement of 630
 - mechanical 627ff
 - natural frequency of 627, 629
 - period of 62, 629
 - resonant 631ff
 - steady-state 633, 634
- Vieta's theorem 609
- Volume of a solid 348ff, 437ff
- Wedge 371
- Work of a variable force 294f, 482ff
- Wronskian 623, 655
- Zero (root) of a function 38
- Zero (trivial) solution of a homogeneous linear differential equation 604

TO THE READER

Mir Publishers would be grateful for your comments on the content, translation and design of this book.

We would also be pleased to receive any other suggestions you may wish to make.

Our address is:

Mir Publishers
2 Pervy Rizhsky Pereulok
I-110 GSP, Moscow, 129820
USSR

Printed in the Union of Soviet Socialist Republics

**OTHER TITLES OF INTEREST
FOR
ENGINEERING AND UNIVERSITY STUDENTS**

**INTRODUCTORY MATHEMATICS FOR ENGINEERS
(LECTURES IN HIGHER MATHEMATICS)**

by A. D. Myskís, D. Sc.

A textbook for students, revised for translation from the 3rd Russian edition. Appears in English for the second time.

**ADVANCED MATHEMATICS FOR ENGINEERS
SPECIAL COURSES**

by A. D. Myskís, D. Sc.

A manual on special courses in mathematics for engineering faculties, and a logical continuation of the author's *Lectures in Higher Mathematics* (already issued in English by Mir Publishers in 1972 and being reprinted in 1975). Presents the subject from the standpoint of modern applied mathematics, with maximum use of intuition and analogy; pays special attention to both the qualitative and the quantitative description of facts. Designed for engineering students, teachers, engineers, and research workers in the applied sciences. Provides a useful bibliography.

EQUATIONS OF MATHEMATICAL PHYSICS. *A. BITSADZE*. Covers traditional sections of the theory of linear integral equations with partial derivatives of second order for elliptical, hyperbolic and parabolic types of elements in the theory of linear integral equations.

MULTIPLE INTEGRALS, FIELD THEORY AND SERIES. AN ADVANCED COURSE IN HIGHER MATHEMATICS. *B. BUDAK AND S. FOMIN*. A concise presentation of branches of mathematics increasingly required by physicists, such as multiple, line and improper integrals, the theory of fields, and power and trigonometric series.

INTRODUCTORY MATHEMATICS FOR ENGINEERS.

***A. MYSKIS*.** An easy-to-follow presentation of higher mathematics with special emphasis on mathematical techniques and practical applications.

PROBLEMS IN MATHEMATICAL ANALYSIS. *B. DEMIDOVICH* (Ed.). Contains 3000 problems covering all branches of higher mathematics (with the exception of analytical geometry) given in courses for engineering students.

PROBLEMS IN HIGHER MATHEMATICS FOR ENGINEERS. *V. MINORSKY*. A collection of 2570 problems in analytical geometry and mathematical analysis with a brief overview of theories, answers and hints that can be used either under the supervision of an instructor or for independent study by correspondence and extramural students.